EXHIBIT 2

Current Perspectives

A Review of Neurofeedback Treatment for Pediatric ADHD

Nicholas Lofthouse[1], L. Eugene Arnold[1], Sarah Hersch[1], Elizabeth Hurt[1], and Roger DeBeus[2]

Abstract

Objective: The aim of this paper was to review all randomized published trials and unpublished conference presentations on the neurofeedback (NF) treatment of pediatric ADHD, and their relevance, strengths, and limitations. **Method:** Via PsychInfo and Medline searches and contacts with NF researchers 14 studies were identified and reviewed. **Results:** The majority were conducted from 1994 to 2010, with 5- to 15-year-olds, usually male and White with the combined type of ADHD. Most studies used theta/beta NF with a unipolar-electrode placement at Cz and demonstrated, where reported, an overall ADHD mean effect size of $d = 0.69$, a medium effect. Main study strengths, within some studies, include use of randomization, treatment control conditions, *Diagnostic and Statistical Manual of Mental Disorders* criteria, evidence-based assessment of ADHD, standard treatment outcome measures, multidomain assessment, and, for some studies, moderate sample size, some type of blind and the identification of medication as a concomitant treatment. Main study limitations (and directions for future research) include the lack of adequate blinding of participants, raters and NF trainers, a sham-NF/blinded control treatment condition, posttreatment follow-up, generalizability, specific details about delivery of NF, identification and control of comorbidity, and the identification, measurement, and control of concomitant treatments and potential side effects. **Conclusion:** Based on the results and methodologies of published studies, this review concludes that NF for pediatric ADHD can be currently considered as "probably efficacious." (*J. of Att. Dis. 2012; 16(5) 351-372*)

In this review, after an introduction to ADHD, we will discuss established treatments (Txs) for ADHD and the need for additional Txs, the concept of neurofeedback (NF), and neurophysiological research on ADHD. Our review of published and unpublished trials (i.e., conference presentations) on the NF Tx of pediatric ADHD will include a description of each study's methods, results, and conclusions followed by a discussion of its relevance to the field and its strengths and limitations. After a summary of reviewed findings, strengths, and limitations, we offer directions for future research and rate the efficacy of NF for pediatric ADHD using American Psychological Association (APA) guidelines (Chambless et al., 1998).

Background and Significance of ADHD

ADHD is characterized by symptoms of inattention, distractibility, overactivity, and impulsivity excessive for developmental age, beginning by age 7, causing impairment in more than one setting, and not better explained by another disorder (*Diagnostic and Statistical Manual of Mental Disorders* [4th ed., Text rev.; *DSM-IV-TR*]; American

Psychiatric Association, 2000). In essence, it involves impairment of the ability to "plan your work and work your plan," while "putting on the brakes." ADHD is the most common childhood mental disorder, with a conservatively estimated prevalence of up to 7% (e.g., Woodruff et al., 2004). It is also chronically impairing, with 60% to 85% of children diagnosed with ADHD continuing to display symptoms into adolescence (Pliszka, 2007) and between 2% and 46% (self-, parent-report, respectively) into adulthood (e.g., Barkley, Fischer, Smallish, & Fletcher, 2002). The core impairments in attention, impulse control, motor modulation, and other executive functions result in more underachievement, discipline problems, impaired social skills, school dropouts, accidents, substance use, occupational

[1]The Ohio State University, Columbus, USA
[2]East Tennessee State University, Johnson City, USA

Corresponding Author:
Nicholas Lofthouse, Department of Psychiatry, The Ohio State University, 1670 Upham Drive, Room 560-A, Columbus, OH 43210-1250, USA
Email: Nicholas.Lofthouse@osumc.edu

inconsistency, and unstable marriages, compared with controls (e.g., Arnold, 2004). Most other mental disorders in childhood and adolescence have a high rate of complication by ADHD comorbidity, and Tx often requires treating ADHD as well as the other disorder (e.g., oppositional defiant disorder [ODD] or conduct disorder = 50%, any anxiety disorder = 25%-30%, learning disorders = 20%-25%, mood disorders = 10%-30%, Tourettes' disorder = 2%; Spetie & Arnold, 2007). Etiology is not firmly established, but the preponderance of evidence suggests that ADHD is a highly heritable disorder (heritability estimate = 76%; Faraone et al., 2005), with abnormalities of neuroanatomy (reduced cortical white and gray matter volume; Sowell et al., 2003), neurochemistry (dopamine and serotonin; Faraone et al., 2005), and neurophysiology (dysfunction of fronto-striatal structures, Bush, Valera, & Seidman, 2005; electrophysiological underarousal of central nervous system, Barry, Clarke, & Johnstone, 2003).

Established Txs for ADHD and Need for Additional Txs

The best documented, most successful, and most widely used Tx for ADHD is catecholaminergic (mainly dopaminergic) stimulant medication (methylphenidate and amphetamine). These show a robust effect in group data, with placebo-controlled effect sizes (ESs; Cohen's d, Cohen, 1988) from 0.7 to 1.5 on parent and teacher ratings of attention and behavior (e.g., Arnold, 2004). Another Food and Drug Administration (FDA)-approved Tx for ADHD is atomoxetine, a noradrenergic reuptake inhibitor, for which one study reported an ES of 0.7 (Michelson et al., 2002). Despite the impressive group data, the response rate at the individual patient level is often less than completely satisfactory. Many of those usually counted as responders in the typically quoted response rate of 66% to 75% have considerable room for improvement.

A better grasp of the actual clinically satisfactory success rate for stimulants can be gleaned from the 579-participant National Institute of Mental Health Multimodal Treatment Study of Children With ADHD (MTA; MTA Cooperative Group, 1999). This study randomly assigned rigorously diagnosed 7- to 9-year-olds with combined type ADHD to one of four Tx conditions, one of which was carefully crafted, intensive medication management (MedMgt) by study staff and another of which was community-care comparison (CC), 66% of whom obtained stimulant medication from their community physician. Swanson et al. (2001) reanalyzed the MTA data to determine the percentage with satisfactory (near-normal) outcome, defined as a mean rating ≤1.0 on ADHD symptoms on a 0 to 3 metric, after 14 months of Tx. Even with the carefully crafted MTA MedMgt, the success rate was only 56%, and for the CC group only 25%. Thus, a conservative estimate of the proportion for

whom stimulant medication by itself is less than completely satisfactory as currently used would be almost half.

A final problem with pharmacotherapy is that an unknown proportion of families refuse to try the approved medications, even though some of their children might benefit. Results from a 2003-2004 national survey of 100,000 U.S. 4- to 17-year-olds estimated that the lifetime prevalence of ADHD was 7.8% with only 4.3% (55% of those with the diagnosis) ever treated with medication for this condition (Centers for Disease Control and Prevention, 2005). This proportion is likely to increase with the negative publicity about these FDA-approved drugs and their boxed warning about cardiovascular effects.

Another established Tx for ADHD is behavioral Tx, which is less effective than well-managed medication (MTA Cooperative Group, 1999) but can sometimes help those who cannot tolerate, fail to respond to, or refuse to try stimulants or atomoxetine. In the MTA, the addition of carefully crafted, intensive behavioral Tx to well-managed medication boosted the percentage with near-normalization (symptom rating of ≤1.0 on 0-3 metric) to 68%. It also reduced the amount of medication required for optimal effect by 20%, thus reducing the risk of side effects. However, even this premium combination Tx, with its intensive MTA behavioral component not available in most communities, still left almost 33% with less than completely satisfactory results. A third of the ADHD population constitutes a considerable public health issue.

In sum, almost a third of children with ADHD do not fully benefit from optimal established Txs and an unknown proportion will not even consider the most effective standard Tx (medication). Therefore, additional complementary and/or alternative interventions are greatly needed.

Background on NF

NF, formerly called electroencephalographic biofeedback and occasionally referred to as neurotherapy, extends to the training of the brain's electrical activity the biofeedback concepts, strategies, and techniques previously established as useful in some medical disorders, especially cardiovascular. Both bio- and NF are thought to work via the classical/operant conditioning mechanisms of learning that train the body/brain to improve its regulation of itself by providing it with real-time video/audio/tactile information about its electrical activity measured from electrodes placed on the surface of the body (biofeedback) or head (NF). For NF, electrodes are placed using the International 10-20 System incorporating a code of letters and numbers to identify the lobe and hemisphere location, respectively (e.g., Cz, a common placement for NF for ADHD, refers to the site over the central lobe [C] at the top of the head [Z = Mid-line]).

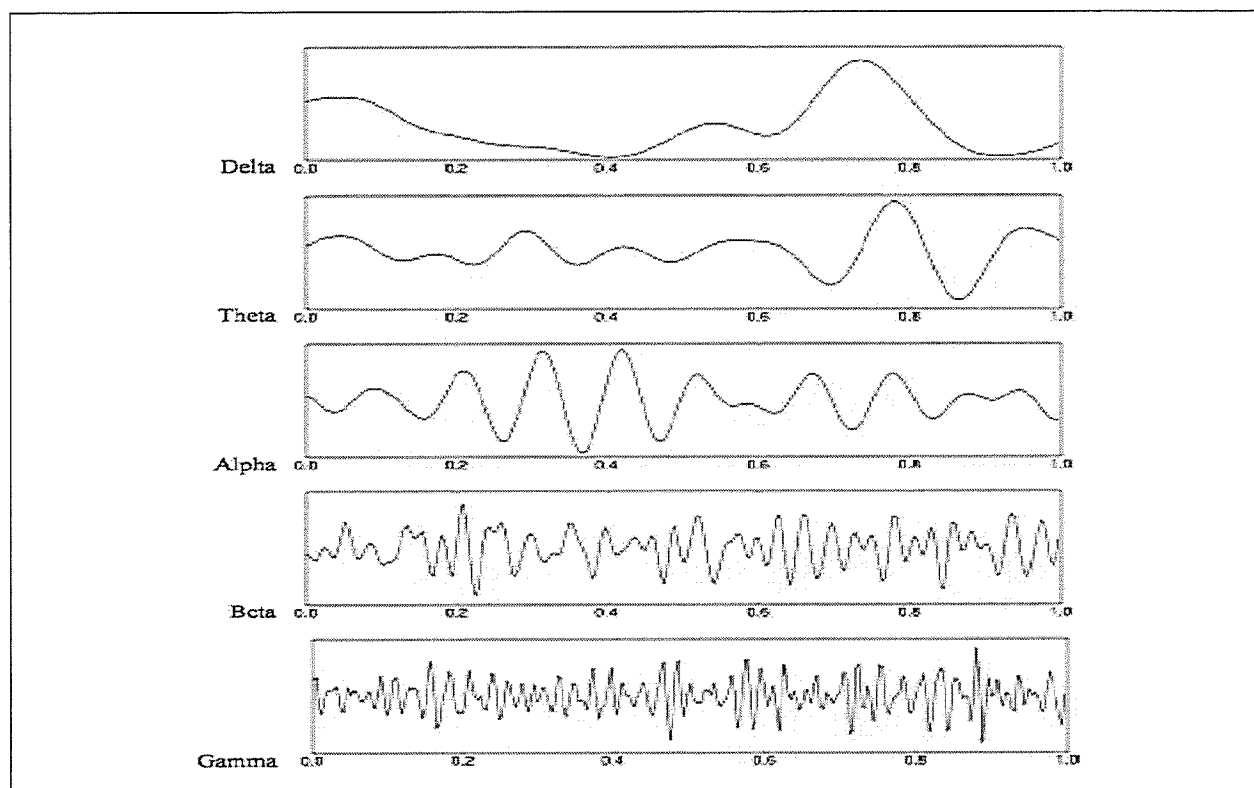First measured in 1924 by Austrian psychiatrist Hans Berger, the brain's electrical activity, shown as brain waves

Lofthouse et al.   353

Figure 1. EEG bands by hertz
Source: Electroencephalography, Wikipedia (2010).

on the electroencephalogram (EEG), was hypothesized to change according to the functional state of the brain while awake or asleep, or in brain diseases such as epilepsy (Berger, 1929). EEG is described in terms of rhythmic activity measured in hertz (Hz), the number of waves per second. It is now known that most of the electrical activity from scalp EEG falls in the range of 1 to 20 Hz. EEG activity is typically divided into specifically named frequency bands (see Figure 1). Up to 4 Hz is called delta (e.g., slow-wave sleep state), 4 to 8 Hz is theta (e.g., drowsy/inattentive state), 8 to 12 Hz is alpha (e.g., relaxed/wakeful state), and 12 to 30 Hz is beta (e.g., active/attentive state). A specific type of low beta activity (12-15 Hz) seen over the sensorimotor cortex, particularly relevant to ADHD, is sensorimotor rhythm (SMR). SMR amplitude is higher when the corresponding sensory-motor areas are idle (e.g., during states of immobility) and decreases when corresponding sensory-motor areas are activated (e.g., during motor tasks). In this manner, SMR is a measure of motor inhibition, strongest when the "brake is on" and weakest when the "brake is off" in these areas.

In addition to these global spontaneous rhythmic activities, other more specific wave patterns (event-related potentials [ERPs]) can be seen in the EEG. ERPs are electrical

representations of underlying sensory and cognitive processing occurring in the brain in response to a stimuli or event (Barry et al., 2003). One specific group of ERPs are slow cortical potentials (SCPs), which were first observed by Walter, Cooper, Aldridge, McCallum, and Winter (1964). SCPs are slow event-related direct-current shifts of the EEG that reflect the excitation threshold of large cortical cell assemblies (Leins et al., 2007; Strehl et al., 2006). Shifts in the negative direction, called the contingent negative variation (CNV), indicate a reduction of the excitation threshold and are thought to be related to cognitive preparation and increased cortical activation of a network, whereas shifts in the positive direction reflect an increase of the excitation threshold and a corresponding inhibition of activation (Drechsler et al., 2007; Leins et al., 2007).

It was first demonstrated that human EEG could be classically conditioned in the mid-1930's (Durup & Fessard, 1935; Loomis, Harvey & Hobart, 1936). However, the actual training of human EEG was not attempted until 1958 when Joe Kamiya, a psychologist at the University of Chicago, began experimenting with teaching an adult to alter his brain-wave frequencies, which were previously thought to be beyond voluntary control. Kamiya's work and NF itself was later popularized by a 1968 *Psychology Today* article

(Kamiya, 1968). In 1973, Nall conducted the first outcome study of NF on 48 children with hyperactivity and learning disorders but found no significant academic or behavioral differences between a group receiving alpha NF and a no-Tx control group. However, three years later, Lubar and Shouse (1976) became the first to report on EEG and behavioral changes in a hyperkinetic child following theta/beta NF, which targeted the reduction of theta waves associated with an inattentive state and increased beta waves associated with an attentive state. Initial methods of giving the brain feedback were, like Kamiya's, via an auditory tone or a visual display of the patient's EEG. More engaging methods were later developed by using visual animation (e.g., simple hands-free video games) and "Go-No-Go" approaches, and more recent technological advances have led to interactive methods using off-the-shelf video games.

Since Kamiya's original work, there has been a significant increase in the clinical application of NF to several psychiatric and medical conditions and, to a lesser degree, a dramatic rise, particularly in the 21st century, in the number of published research and dissertation studies (e.g., *PsychInfo/Medline* journal searches for title terms, neurofeedback, electroencephalographic/EEG biofeedback, or neurotherapy: pre-1970 = 9 studies, 1970-1979 = 19 studies, 1980-1989 = 20 studies, 1990-1999 = 14 studies, 2000-2010 = 79 studies). In 1995, the International Society for NF and Research (ISNR; www.isnr .org) and accompanying *Journal of Neurotherapy* were founded, with annual ISNR conferences since 1993. More recently, the field of NF was introduced to the general public by Jim Robbins's (2008) informal and engaging review of the history of the field, *A Symphony in the Brain: The Evolution of Brain Wave Biofeedback.*

Theoretical Basis for NF in the Tx of ADHD

In terms of NF's relevance for ADHD, neuroimaging studies suggest that ADHD is associated with smaller and possibly underaroused frontal lobes and other brain regions responsible for sustained attention and behavioral planning and motor control (Swanson & Castellanos, 2002). Positron emission tomography and single-photon emission computed tomography studies have reported reduced blood flow and metabolism suggesting electrophysiologic underarousal over frontal and central-midline cortical regions in approximately 80% to 90% of patients with ADHD (Chabot, Merkin, Wood, Davenport, & Serfontein, 1996; Clarke, Barry, McCarthy, & Selikowitz, 2001; Mann, Lubar, Zimmerman, Miller, & Muenchen, 1992; Monastra et al., 1999).

Research has also demonstrated that many patients with ADHD have more slow-wave (especially theta, 3.5-8 Hz) power in their resting EEG spectral analysis than normal controls, and conversely less beta (12-20 Hz) power especially in central and frontal regions, most probably reflecting

underarousal of the central nervous system (see review by Barry et al., 2003). The lower beta frequencies (12-15 Hz, SMR) have been found to be associated with calm immobility in experimental animals, whereas the higher beta frequencies (>15 Hz) are associated with focusing on a task or other situation requiring attention. Research on ERPs associated with ADHD has also identified deviant sensory and cognitive processing for early and late stages of the evoked response (see review by Barry et al., 2003). Research on SCPs has shown that children with inattention and hyperactivity have reduced cortical negativity (i.e., a deviant CNV) during cognitive preparation (e.g., Banaschewski et al., 2004; Rockstroh, Elbert, Lutzenberger, & Birbaumer, 1990), suggesting that "failure to engage specific cortical networks contributes to the performance decrement" (Strehl et al., 2006, p. 1532). Theoretically, this slow-wave activity/reduced cortical negativity or underarousal is associated with the core ADHD symptoms of inattention, hyperactivity, and impulsivity.

Furthermore, methylphenidate, a proven Tx for ADHD, affects the EEG frequencies. For example, Song, Shin, Jon, and Ha (2005) reported that during a continuous performance test (CPT), methylphenidate increased alpha in frontal and occipital areas, increased beta in almost all areas, decreased theta in occipital and right temporal–parietal areas, and mildly decreased delta in the occipital–parietal areas. Loo and Barkley (2005) reported that, compared with placebo, methylphenidate increased alpha activity in central and parietal regions, and, in clinical medication responders, it increased the frontal beta activity whereas in clinical non-responders, it decreased frontal beta activity. The increase in frontal beta significantly correlated with improvement on CPT and parent ratings of attention and behavior. Decreased right frontal theta correlated with parent ratings of improved attention. Thus, it appears that changes in brain-wave frequencies are associated with stimulant medication response.

In sum, research on the physiological basis of ADHD, EEG dysfunctions and their relationship to underlying thalamocortical mechanisms, changes associated with a positive medication response, and the idea that brain waves can be consciously learned has formed the theoretical and evidence-based foundation of NF in the Tx of ADHD.

Research on the NF Tx of ADHD

In a recent review of the literature, Monastra (2005) noted that, over the past 25 years, numerous studies have reported benefit from NF in ADHD, leading him to conclude that NF is "probably efficacious" for ADHD, the middle level on a five-level grading of evidence base used by the APA. Hirshberg, Chiu, and Frazier (2005), editors of the special EEG issue of *Child and Adolescent Psychiatric Clinics of North America*, in which Monastra's review appeared, were more enthusiastic. They first stated, "EBF

[i.e., NF] meets the AACAP [American Academy of Child and Adolescent Psychiatry] criteria for 'Clinical Guidelines' for Treatment of ADHD" (p. 12), the third level of four, which is based on limited evidence such as open trials/case studies and/or strong clinical consensus and should always be considered by the clinician. Hirshberg et al. also stated that "the use of EBF [i.e., NF] for ADHD will (with the publication of the second RCT) meet the most stringent APA criterion of efficacious and specific, which requires two independent RCT's, among other factors" (p. 13). The second randomized controlled trial (RCT; the first by Linden, Habib, & Radojevic, 1996) refers to the Levesque, Beauregard, and Mensour study published in a peer-reviewed journal in 2006. Therefore, in contrast to Monastra, Hirshberg et al. believed that, by APA's criteria, NF is now an "efficacious treatment."

Monastra's (2005) review identified four controlled (i.e., with a non-NF comparison group) studies of NF for pediatric ADHD (Fuchs, Birbaumer, Lutzenberger, Gruzelier, & Kaiser, 2003; Linden et al., 1996; Monastra, Monastra, & George, 2002; Rossiter & La Vaque, 1995), with only Linden's study involving randomization of participants to groups. Until their publication, the literature contained only case studies and open or nonrandomized trials. This is problematic as without randomization, results are difficult to interpret as they may in fact be due to the specific effects of NF; or due to selection effects and associated expectations because of participants/parents choosing their preferred Tx group, nonrandom participant experiences (i.e., participant history), regression to the mean, maturation, or practice with assessment measures; or due to the interaction of any of the above factors. Due to these problems with nonrandomized studies, the remainder of this review will only consider randomized studies.

Another threat to the internal validity of NF research is that of nonspecific Tx effects defined as "outcomes that are a result of intervention components or processes that are not currently specified in the intervention theory" (Donovan, Kwekkeboom, Rosenzweig, & Ward, 2009, p. 986). Nonspecific Tx factors can include attention by the experimenter/trainers, practice paying attention/sitting still/inhibiting responses, provider qualities, Tx structures, therapeutic alliance, participants' motivation for improvement, and participations'/ experimenters'/trainers'/raters' (i.e., parents, teachers, and clinicians) expectations (positive and negative) regarding study outcome (Donovan et al., 2009). People's expectations about Tx outcome are ideally controlled in Tx outcome research by methodology preventing participants/experimenters/ trainers/raters from knowing which condition the participant has been randomly assigned (so-called blinding). Other nonspecific Tx effects can be controlled by using a "fake" Tx (sham/placebo) Tx or another Tx that is practically identical to the target Tx except for the specific active Tx component, in this case feedback contingent on the person's EEG.

For example, a sham-NF condition might give inaccurate feedback not contingent on the person's EEG. Without Tx control groups matched in duration, intensity, and apparatus and blinding of participants/experimenters/raters, it is impossible to separate specific effects from nonspecific effects of Tx. Only a few NF studies have used such features as sham NF to control for nonspecific Tx effects.

Since Monastra's (2005) review of nonrandomized and randomized published studies, an additional seven studies (identified by September 2010 *PsychInfo* and *Medline* searches and contacts with researchers in the NF field) have been published using randomization (Levesque et al., 2006; Drechsler et al., 2007; Leins et al., 2007; Gevensleben, et al 2009a; Gevensleben et al 2009b; Holtmann et al., 2009; Perreau-Linck, Lessard, Levesque, & Beauregard, 2010; deBeus & Kaiser, 2011) and one meta-analysis of randomized and nonrandomized published and unpublished trials (Arns, de Ridder, Strehl, Breteler, & Coenen, 2009). As Drechsler et al.'s (2007) randomization was incomplete and without any indication of how many participants were correctly randomized, their study is not considered further. In addition, deBeus and Kaiser's study was a reanalysis of their 2006 conference presentation and is henceforth considered as one study rather than two studies. In addition to these published studies, via references within the preceding publications, the ISNR website (www.isnr.org), and contacts with researchers in the NF field, eight unpublished studies presented at conferences were identified (Fine, Goldman, & Sandford, 1994; Palsson et al., 2001; Orlandi & Greco, 2004; deBeus, 2006; Picard, Moreau, Guay, & Achim, 2006 [included two randomized studies]; McGrady, Prodente, Fine, & Donlin, 2007; Urichuk et al., 2009). Although unpublished studies suffer from not meeting the same scientific standards of peer-review, we have included them to examine the research on this topic as thoroughly as possible, document information that may be eventually lost over time and because conference presentations have inserted them into the public nowledge base via their presentations at conferences. Authors of the unpublished studies were contacted via email on several occasions to provide specific information regarding their studies, but where information was not provided or available, "data unavailable" is noted. Therefore, to date, there have been 14 randomized studies of NF on youth with ADHD, 6 published and 8 unpublished (including 2 studies by Picard et al. [2006]).

The majority of these studies used traditional NF, which typically induced beta-band EEG rhythms, either SMR (12-15 Hz) or higher (15-18 Hz), and suppressed theta rhythms (4-8 Hz), by visual and auditory feedback (subsequently referred to as theta/beta and, when specified, SMR). More recently, some researchers have examined NF of SCPs, which involves training children to increase their cortical negativity to mobilize attentional resources and self-regulatory

　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　Journal of Attention Disorders 16(5)

capacities (Strehl et al., 2006). In 2004, Heinrich, Gevensleben, Freisleder, Moll, and Rothenberger were the first to successfully use SCP NF to change the polarity of the EEGs (i.e., positivity vs. negativity) of a group of children with ADHD. Although Heinrich and colleagues did not use randomization, there have since been three studies of SCP NF with pediatric ADHD using randomization (Drechsler et al., 2007; Gevensleben, et al., 2009b; Gevensleben, 2009a; Leins et al., 2007), with the latter using a control group whereas the other two used a theta/beta NF comparison group.

Following, in chronological order, are summaries and critiques of the 14 randomized studies of NF on pediatric ADHD. Due to the limitations of null-hypothesis significance testing for interpreting social science data (Ferguson, 2009), we follow the recommendations of the Wilkinson Task Force (Wilkinson and Task Force on Statistical Inference, 1999) and, wherever possible, calculate and report ESs for each reported marginally significant and significant result using Cohen's d (Cohen, 1988). Assuming randomization controls for any pre-Tx differences between groups, we based our ES estimate on post-Tx means (M) and standard deviations (SD), where given, of the Tx and control groups. For reference, ds between 0.2 and 0.4 are considered small ESs, 0.5 and 0.8 medium, and >0.8 large.

1. **Fine et al (1994)** presented the very first documented randomized study of NF for pediatric ADHD at the 1994 APA conference. Their sample included seventy-one 8- to 11-year-olds from the United States (85% male, ethnicity data unavailable), with ADHD diagnosed by a physician or psychologist (data on whether ADHD was via *DSM* criteria, and ADHD type and comorbidity unavailable). It was reported that the "majority" were taking medication, and pre–post Tx assessment was given prior to daily administration of medicine (information on other concomitant Txs or change in any concomitant Txs during study was unavailable). Participants were randomized to theta/beta/SMR (unipolar-electrode placement) NF via a video game, cognitive training (Captain's Log program), or no Tx with 23 to 24 in each condition. NF consisted of twenty 30-min, twice-per-week sessions over a 10-week period. The cognitive-training condition also involved twenty 30-min, twice-per-week Tx sessions over a 10-week period. Children, parents, and examiners were not blind to group affiliation. Compared with the no-Tx group, both the NF and cognitive-training groups showed significant improvement on 8/21 parent-rating scales and 4/30 laboratory test variables. These included, for parent scales, the Revised Conners's Questionnaire Learning Problems, Impulsive-

Hyperactivity, and Hyperactivity Index Scales (all $p < .01$, Goyette, Conners, & Ulrich, 1978); Child Behavior Checklist (CBCL) Schizoid/Anxious, Depressed, Social Withdrawal and Hyperactive Scales (all $p < .01$, Achenbach, 1991); and the Home Situations Questionnaire M severity score ($p < .05$, Altepeter & Breen, 1989) and, for the laboratory tests, the Wide Range Assessment of Memory and Learning ($p < .01$, Sheslow & Adams, 1990); Number/Letter Memory (measure of attention, $p < .01$); Design Memory (visual memory, $p < .01$) and Story Memory (verbal memory, $p < .01$); and Stroop Test (inhibition, $p < .005$; Stroop, 1935). M and SD were unavailable, so ESs could not be calculated.

CRITIQUE: Being the first RCT of NF for pediatric ADHD, this was a very important study for the field and more convincing than all previous uncontrolled studies because of randomization (to control for selection effects, participant history, regression to the mean, maturation, and practice with assessment measures), a Tx control condition (to control for nonspecific Tx effects), multidomain assessments, and standard Tx outcome measures used in ADHD trials. Unfortunately, the use of a cognitive-training control group to teach attention skills may have made it more challenging to obtain significant comparative results for the NF group. Although two of the three core symptom clusters of ADHD (i.e., hyperactivity and impulsivity) improved relative to the untreated control condition, they did not differ significantly from the comparison Tx. Study limitations included the few significant results for both Tx groups (12 of 51 variables), the fewer observed improvements for the NF group alone ($n = 6$), and the lack of follow-up (FU) data to examine long-term effects. In addition, as participants, parents, and experimenters were not blind to study condition, participant-, parent-, and experimenter-expectancy effects were not controlled. Finally, because comorbidity and the presence of and change in any concomitant Txs were not identified, these confounds pose additional potential threats to internal validity.

2. **Linden et al (1996)** published the first randomized study of NF for pediatric ADHD They randomly assigned 18 U. S. children (gender and ethnicity not reported), aged 5 to 15, diagnosed with *DSM-III-R* (American Psychiatric Association, 1987) ADHD or ADD (33% learning disorder comorbidity) to a wait-list control (WLC) condition or forty 45-min, twice-per-week, theta/beta (bipolar Cz and Pz [Posterior Lobe at Mid-line Z]) NF sessions over 6 months. Children, parents, and examiners were not blind to group affiliation. No participants were

on any medication for ADHD or involved in any other Txs during the study. The NF group demonstrated a significant ($p = .02$) increase (9 IQ points) in composite IQ (Kaufman-Brief Intelligence Test; Kaufman & Kaufman, 1990) and a significant ($p = .04$) reduction in parent-rated inattention (Swanson, Nolan, and Pelham [SNAP] Rating Scale; Swanson, Nolan & Pelham, 1981) compared with the wait-list group. However, there were no significant differences between groups on parent ratings of hyperactivity-impulsivity or aggressive behavior. *M* and *SD* were unavailable, so ESs could not be calculated.

CRITIQUE: As the first published randomized study of NF for pediatric ADHD, this study was a crucial step forward for the field. Study strengths included randomization, an untreated control group, *DSM* diagnoses, multidomain assessments, standard Tx outcome measures, and control of concomitant Tx. Study limitations included the lack of participant-, parent-, and experimenter blinding, a Tx or sham-control condition, and FU data. In addition, although the core ADHD symptoms of inattention showed significant results, hyperactivity-impulsivity did not.

3. **Palsson et al. (2001)** compared standard NF with a new video game NF and presented their results at the annual meeting of the Association for Applied Psychophysiology and Biofeedback (AAPB). A total of 22 (86% boys, ethnicity data not collected) American 9- to 13-year-olds with *DSM-IV* (APA, 1994) ADHD hyperactive-impulsive type (via *DSM* criteria and physician diagnosis) with no history of affective problems, mental retardation, or learning disorders were randomized to either standard NF ($n = 11$) or Sony PlayStation video game NF ($n = 11$). All children were on short-acting ADHD medication, but any change in medication and the presence of and change in other concomitant Txs was not reported. Both types of NF were designed to reduce theta/alpha and strengthen SMR/beta (unipolar Cz) and were given for 60 min, once-twice per week, for 40 sessions for approximately 20 weeks. There was no blinding. Scores on the Behavior Assessment System for Children–Monitor (BASC, Reynolds & Kamphaus, 1992) and Test of Variables of Attention (TOVA; Leark, Greenberg, Kindschi, Dupuy, & Hughes, 2007) indicated similar significant pre–post improvements for both groups. On the BASC, inattention scores significantly decreased for video game NF ($p = .002$) and standard NF ($p = .001$), hyperactivity scores significantly decreased for video game NF and standard NF (both $p = .02$),

and internalizing scores significantly decreased for video game NF ($p = .01$) and standard NF ($p = .005$). On the TOVA, video game NF and standard NF had significant/marginally significant improvements for errors of omission, a measure of inattention ($p < .09$ and $p < .05$, respectively); errors of commission, a measure of impulsivity (both $p < .02$); total number of correct responses ($p < .05$ and $p < .03$, respectively); and *D* prime, a measure of the ability to discriminate target stimuli from nontarget stimuli ($p < .004$ and $p < .01$, respectively). As the authors noted, these results are hard to interpret because the video game–NF group was superior to the standard-NF group on their measures at baseline. Parents' subjective appraisal of Tx effect on their child's ADHD was marginally significant ($p < .07$) with more positive ratings for video game NF (65%) than for standard NF (48%). Video game NF was also rated significantly higher than standard NF on a 10-point enjoyability rating by parents (9.36 vs. 6.75, $p < .03$) and children (9.5 vs. 7.8, $p = .03$). Trends on pre–post quantitative EEG (QEEG) change maps indicate that video game NF may have more positive effects than standard NF. Finally, both types of NF were reported to improve children's functioning substantially (p not available) above the benefits of background medication. *M* and *SD* were unavailable, so ESs could not be calculated.

CRITIQUE: This study is important as it introduced a new NF technology using off-the-shelf video games to try to make NF more interactive, engaging, and familiar to children. This was supported by the results showing that parents and children enjoyed the video game–NF modality more, parents considered it as more effective, and it was associated with more positive QEEG changes than standard NF. Study strengths included randomization, *DSM* diagnoses, control of comorbidity, multidomain assessments, standard Tx measures, satisfaction ratings, and an alternative NF condition. This study also found significant pre–post improvements for the core symptoms of ADHD. Study limitations included the absence of a sham-NF condition, blinding and FU data, and no statistical between-group comparisons of the two Txs, only within-group comparisons.

4. **Orlandi and Greco (2004)** presented the results of a randomized study at the annual meeting of ISNR. Their sample included thirty-six 9- to 11-year-old American boys (86% White) with *DSM-IV* ADHD-combined type without comorbidity, medication, or psychotherapy (change in educational services was not monitored). Children

and Tx staff were not blind, but evaluations were done by blinded clinicians and blinded parents. In all, 17 children were randomized to forty, 45-min theta/beta/SMR (unipolar Cz) NF sessions twice per week for 20 weeks and 19 children to a control condition of equal duration, frequency, and intensity in the same setting. The control activity was playing a video game designed to improve attention, visual tracking, hand–eye coordination, attention to detail, planning, concentration, memory, and patience. The NF group, but not the control group, showed significant ($p < .05$, $d = 0.98$) improvement on pre–post Tx parent ratings (item sum from Conners's Parent Rating Scale–Revised [CPRS-R]; Conners, 2002a) and significant ($p < .01$, $d = 1.6$) improvements on independent blinded-clinician ratings of symptom severity (Clinical Global Impression–Severity; Guy, 1976).

CRITIQUE: This study advanced the field by being the first randomized study of NF for pediatric ADHD to use blinding. Additional strengths included standard Tx outcome measures, *DSM* diagnoses, control of comorbidity, medication and psychotherapy, and a control condition. The use of a cognitive-training video game control group may have made it more challenging to obtain significant comparative results for the NF group. Study limitations included the absence of a blind for children and Tx staff, sham-NF condition, and FU data. The two Tx groups did not actually separate significantly (a possible power problem) on the critical blinded-parent ratings; it was just that NF showed significant improvement and the control improvement failed to reach significance. The failure of the control group to reach significance could also be a power problem with such a small sample. As Orlandi and Greco (2004) reported a significant global change for CPRS ADHD items, it is difficult to know whether there were significant changes for all three core ADHD symptoms of inattention, hyperactivity, and impulsivity. Finally, although educational services continued, because they were not monitored, it is not known whether those services changed during the study and were a possible confound.

5. **Levesque et al. (2006)** published a study of randomly assigned fifteen 8- to 12-year-olds (80% male; ethnicity not reported) with *DSM-IV* ADHD (ADHD type not reported; no comorbidity). NF involved forty 60-min theta/beta/SMR (unipolar Cz) sessions, 3 times per week, over 13½ weeks. Five additional participants were randomized to a wait-list condition. No participants took psychostimulants, but the presence of and change in other concomitant interventions during the study were not reported. Blinding of participants,

experimenters, and raters was also not reported. Those receiving NF, but not those on the wait-list, improved significantly ($p < .05$; $d = 0.54$) on the Digit Span subtest of the Wechsler Intelligence Scale for Children–Revised (WISC-R; Wechsler, 1974); Integrated Visual and Auditory (IVA; Sandford & Turner, 1995) CPTs ($p < .005$, $d = 0.37$); CPRS-R (Conners, 2002a) scales of inattention ($p < .001$, $d = 1.62$) and hyperactivity ($p < .05$, $d = 0.98$); Counting Stroop Task Neutral ($p < .05$, $d = 0.85$) and interference ($p < .05$, $d = 1.02$) trials (Bush et al., 1998); and, reported in Beauregard and Levesque (2006), Go/No-Go task ($p < .005$, $d = 0.63$) and No-Go trials ($p < .005$, $d = 1.55$). Most importantly, only the NF group had significant (p not given) pre- to post-Tx functional magnetic resonance imaging (fMRI) activation in the right anterior cingulate cortex (ACC), left caudate nucleus, and left substantia nigra during the Counting Stroop Task and in the right ventrolateral prefrontal cortex, right dorsal ACC, left thalamus, left caudate nucleus, and left substantia nigra during the Go/No-Go task (Beauregard & Levesque, 2006). This suggested to Levesque and colleagues that, for children with ADHD, NF "has the capacity to functionally normalize the brain systems mediating selective attention and response inhibition" (Beauregard & Levesque, p. 3).

CRITIQUE: Although a small sample, this study was important because it was the first to show pre–post Tx neurophysiological changes in brain areas reported to be smaller/hypoactive in ADHD by numerous reputable investigators. Additional study strengths included randomization, a control group, control of comorbidity, *DSM* diagnoses, multidomain assessment, standard Tx outcome measures, and significant results for objective measures and parent ratings of the core ADHD symptoms of inattention, hyperactivity, and impulsivity. It is also notable that this study used thrice-weekly Tx in contrast to most previous studies, which used twice weekly Tx, which suggests that more frequent NF may be a viable option to speed up improvement over time. Unfortunately, there was no blinding of participants, raters, and experimenters; a sham-NF condition; or FU data. However, it seems unlikely (though possible) that expectancy biases would activate those particular brain areas in all those receiving NF. Finally, Levesque et al.'s (2006) statistical comparisons are questionable as they did not use the standard statistical method of directly comparing change scores between two groups (i.e., a 2 × 2 ANOVA of the NF and WLC groups' pre- and post-Tx scores). Instead, they initially used a between-group *t* test to compare the two groups' pre-Tx scores and, after finding no significant differences, used separate within-group *t* tests to examine pre- to postchanges

Lofthouse et al. 359

separately for the NF and WLC, which is not a direct or valid comparison and effectively neutralizes the use of a WLC.

6. **deBeus (2006)** presented results from a triple-blind (i.e., children, raters [parents and teachers], and NF trainers) crossover randomized controlled design at the American Psychiatric Association annual meeting. He examined 53 American 7- to 11-year-olds (66% males, 91% White) diagnosed with *DSM-IV* ADHD (47% inattentive type, 53% combined; 34% comorbid with ODD, 13% learning disorder) via the Diagnostic Interview for Children and Adolescents-IV (DICA-IV; Reich, Leacock, & Shanfeld, 1997) More than half (57%) were on medication (presence of concomitant psychotherapy and educational interventions and change of concomitant Tx during the study were not monitored). Two days prior to screening and during pre- and post-Tx assessment, participants were removed from any prescribed stimulants. In a randomly assigned order, participants received twenty 30-min theta/beta/SMR (unipolar Fz [frontal lobe, midline]) NF sessions, over 10 weeks and 20 sham-feedback sessions of equal frequency, duration, and intensity in the same setting, for a total of 40 Tx sessions. The sham-NF condition involved random rewards from the same equipment to control for nonspecific Tx effects. However, another technician who checked and adjusted the feedback was not blind but was separated from and did not have any contact with the child and NF trainer. This presentation included participants who had completed the first 20 sessions to which they were randomly assigned (i.e., either real or sham NF). Compared with the sham sessions, participants receiving NF had significantly better response control ($p < .0001$) and attention ($p < .0004$) on the IVA (Sandford & Turner, 1995), and significantly lower ($p = .0002$) parent-rated *DSM-IV* inattentive symptoms (CPRS-R; Conners, 2002a) and teacher-rated *DSM-IV* inattentive ($p = .05$) and hyperactive-impulsive ($p < .0001$) symptoms (Conners's Teacher Rating Scale–Revised [CTRS-R]; Conners, 2002b). *M* and *SD* were unavailable, so ESs could not be calculated.

A reanalysis of this data was recently published by **deBeus and Kaiser (2011)** including only the participants who completed all 40 Tx sessions of real and sham NF (42/53). This reanalysis involved 42 American 7- to 11-year-olds (31% males, 90% White) diagnosed with *DSM-IV* ADHD (43% inattentive type, 57% combined; 40% comorbid with ODD, 33% anxiety spectrum, and 17% dysthymic disorders), 57% of whom were on medication. After all 40

sessions, "NF learners" (NF-L, that is, children whose beta/theta + alpha [Engagement Index] EEG improved 1 *SD* pre–post active-NF Tx, $n = 31$) and "NF nonlearners" (NF-NL, that is, children who did not meet criterion, $n = 11$) were identified. Compared with the sham sessions, during active NF, NF-L had significantly better scores on the CTRS-R (Conners, 2002b) ADHD Total ($p < .005$, ES = 0.50), Inattentive ($p = .01$, ES = 0.41), and Hyperactive-Impulsive ($p = .02$, ES = 0.37) Scales and on the IVA (Sandford & Turner, 1995) response control ($p = .003$, ES = 0.63) and attention ($p = .002$, ES = 0.60) tasks. No Tx effects were found for the CPRS-R (Conners, 2002a).

CRITIQUE: Along with the following study by Picard et al. (2006), deBeus' (2006) study was an essential development in the evolution of NF research as it used a sham-NF condition and triple-blind design. Along with the 2010 NF-L data reanalysis, deBeus' study is the most sophisticated to date. Other study strengths include randomization, a reasonable sample size, multidomain assessment, standard Tx outcome measures, *DSM* diagnoses, and measurement of comorbidity. In addition, deBeus found significant results for objective measures and teacher ratings of the core ADHD symptoms of inattention, hyperactivity, and impulsivity. However, instead of a comparison of sham versus active conditions for the NF-L, it would have been statistically more appropriate to run a 2 × 2 analysis comparing the NF-L versus NF-NL in active versus sham. In addition, as is required of all crossover designs, there was no examination of a possible order effect, and there were possible carryover and learning effects (K. Conners, personal communication, February 15, 2011). Other limitations include the absence of a validity check on the blind, presence of or change in concomitant Txs during the study not being monitored and there were no FU data. Although published in an edited book, the details have not undergone peer-reviewed journal publication.

7. **Picard et al. (2006)** presented their results from two randomized studies at the ISNR annual meeting. The first evaluated the effect of NF compared with a WLC to select relevant variables and provide data for the second study where a sham-NF group was added. In the first study, they examined 15 Canadian 7- to 12-year-olds (87% males, 100% White) diagnosed with *DSM-IV* ADHD by a physician and neuropsychologist (100% combined; comorbidity data unavailable), all of whom were on medication and had "standard support programs" (presence of concomitant psychotherapy and educational interventions and change of concomitant Tx during the study were not reported). Children were randomly assigned to wait-list ($n = 7$) or real NF ($n = 8$), which consisted of forty 29- to 42-min theta/SMR NF sessions, 3 times per week, over

13 weeks. Compared with wait-list, participants receiving real NF had significantly improved parent-rated hyperactive ($p = .006$), inattentive ($p = .009$), and global ($p = .002$) scores (DuPaul Behavioural Questionnaire [DBQ]; DuPaul, Power, Anastopoulos, & Reid, 1998) and improved ($p = .018$) Verbal Comprehension Index [VCI] scores (Wechsler Intelligence Scales for Children–4th ed.; Wechsler, 2003).

8. In their second study, **Picard and colleagues (2006)** examined 31 separate children (81% males, 100% White) aged 7 to 12 diagnosed with *DSM-IV* ADHD by a physician and neuropsychologist (100% combined type; comorbidity data unavailable), all of whom were on medication and had "standard support programs." Participants were asked not to start any other Txs during the study, and to the authors' knowledge, none of them did. Children were randomly assigned to wait-list ($n = 11$), real NF ($n = 10$), or sham NF ($n = 10$). The real and wait-list conditions were like those in stage 1, but the sham-NF condition involved random rewards from the same equipment to control for nonspecific Tx effects. Participants, parent raters, and 75% of the NF trainers were blind to condition assignment. Similar to their first study, compared with wait-list, real NF had marginally or significantly improved hyperactive ($p = .068$), inattentive ($p = .019$), and global ($p = .009$) parent-DBQ and WISC-IV scores ($p = .045$). Compared with sham NF, real NF showed marginally or significantly improved hyperactive ($p = .074$), inattentive ($p = .057$), and global ($p = .013$) parent-DBQ and VCI scores ($p = .062$). In contrast, the wait-list and sham-NF groups were not significantly different on parent-DBQ or VCI scores. Picard and colleagues concluded that "these results indicate that the benefits observed after neurofeedback are not the result of motivational and social variables embedded in the treatment" and "the EEG modification component is essential to obtain the NFB (NF) effects." *M* and *SD* were unavailable, so ESs could not be calculated.

CRITIQUE: Similar to the deBeus's (2006) study, Picard et al.'s (2006) use of a sham-NF condition and triple-blind design made a major contribution to the field. Other study strengths included randomization, multidomain assessment, standard Tx outcome measures, *DSM* diagnoses, and significant results for objective measures and ratings of the core ADHD symptoms of inattention and hyperactivity. Unfortunately, there were no validity check of blinding, FU data, and this study has not undergone peer-reviewed publication.

9. **Leins et al (2007)** published the first randomized study of SCP training and the first to compare SCP with theta/beta NF. They randomized thirty-eight German 8- to 13-year-olds (84% male, ethnicity not reported) with *DSM-IV* ADHD (79% combined, 21% inattentive type; 24% comorbidity) to thirty 60-min, five-per-week sessions of either theta/beta (bipolar C3f [central lobe, left hemisphere] and C4f [central lobe, right hemisphere]) or SCP (unipolar Cz) NF training over three Tx phases of 2 weeks training interspersed with 4- to 6-week breaks. Children, parents, and teachers were blind to assignment but not the NF trainer (blindness status of assessment examiner not reported). Most participants (95%) of the sample were nonmedicated, but the presence of other concomitant Txs and change in any concomitant Txs during the study were not reported.

Outcome variables were examined between baseline, post-Tx (post), and at 6-month FU, and the following ESs (*d*) were reported by the authors. No time-by-Tx effects were significant, but the size of changes in the two groups are of interest. On parent ratings of *DSM-IV* ADHD symptoms, significant improvements on attention (baseline-post: $p = .009$, $d = 0.80$; baseline-FU: $p = .026$, $d = 0.78$) and hyperactivity (baseline-post: $p = .006$, $d = 0.34$; baseline-FU: $p = .002$, $d = 0.61$) were found for theta/beta NF. On teacher ratings of *DSM-IV* ADHD symptoms, significant improvements for hyperactivity were reported for both SCP (baseline-FU: $p = .003$, $d = 0.56$) and theta/beta (post-FU: $p = .015$, $d = 0.50$), impulsivity for theta/beta alone (post–FU: $p = .006$, $d = 0.56$), and social behavior for SCP alone (baseline-FU: $p = .015$, $d = 0.59$). Interestingly, hyperactivity and impulsivity of the theta/beta group deteriorated from baseline-post, before improving significantly post-FU. Parent ratings of the frequency of problems at home (Eyberg Child Behavior Inventory; Eyberg & Pincus, 1999) decreased significantly for SCP (baseline-post: $p = .033$, $d = 0.43$; baseline-FU: $p = .048$, $d = 0.45$) but not for theta/beta. CPRS-R (Conners, 2002a) parent ratings showed a significant improvement for theta/beta (baseline-FU: $p = .009$, $d = 1.02$) but not for SCP. Between baseline-FU, both theta/beta and SCP had significant ($p < .001$ and $p = .008$, $d = 0.82$ and 0.54, respectively) increases in performance IQ (Hamburg Wechsler Intelligenztest fur Kinder–Dritte Auflage [HAWIK-III]; Tewes, Rossmann, & Schallberger, 1999) but only theta/beta had significant ($p = .015$) increases in full-scale IQ ($d = 0.62$). Attention, as measured by a German version of a standardized computer battery, significantly improved for both SCP (baseline-post: $p < .001$, $d = 0.92$; baseline-FU: $p < .001$, $d = 1.09$) and theta/beta (baseline-FU: $p = .021$, $d = 0.66$). Leins and colleagues (2007) concluded that "this is the first time that stability of clinical effects after a

neurofeedback treatment is demonstrated six months after the treatment" (p. 86).

CRITIQUE: This study is notable for being the first randomized examination of SCP NF, the first comparison of SCP NF to theta/beta NF, the first to intersperse 4- to 6-week breaks during NF training, and the first to conduct a 6-month post-Tx FU. It also had the strengths of blinding children, parents, and teachers; examining NF 5 times per week, *DSM* diagnoses, measurement of comorbidity, multidomain assessment, standard Tx outcome measures; and finding significant pre–post results for the core ADHD symptoms of parent-rated inattention and hyperactivity and teacher-rated inattention and impulsivity. Unfortunately, there was no blinding of NF trainers (and possibly assessment examiners) and no sham-Tx condition. The authors acknowledged such limitations and noted a double-blind design was not feasible because the trainer had to adjust the feedback parameters and because a placebo condition or wait-list for a Tx of 30 sessions was ethically incompatible with the Declaration of Helsinki (World Medical Association, 2000). As several other NF researchers and practitioners over the years have cited ethical issues with using a placebo condition, we will return to a discussion of this topic later on in the article.

A final criticism of this study is the lack of control or measurement of potential confounds related to changes assessed 6 months after the last NF session, which means that any changes at FU may have been due to NF and/or non-NF factors such as nonrandom participant experiences (e.g., beginning medicine, psychotherapy, or school services); regression to the mean; maturation; practice with assessment measures; participant-, observer-, and experimenter-expectancy effects; or a combination of any of the preceding factors. This is particularly problematic for this study as only 4/17 significant results occurred between baseline and post-Tx (i.e., parent ratings of *DSM-IV* attention and hyperactivity, parent ratings of problems at home, and computerized measures of attention). The remaining 13 significant results occurred between baseline and FU and post-Tx and FU, which could indicate either a "sleeper effect" of NF and/or the post-Tx effect of nonspecific factors. Therefore, this casts serious doubt on their conclusion that "this is the first time that stability of clinical effects after a neurofeedback treatment is demonstrated six months after the treatment" (p. 86). The study results are in fact hard to interpret because the design and results, with no non-NF comparison group and no significant differences between groups, do not allow a conclusion of controlled efficacy for either NF Tx. Although this was a randomized trial, the results essentially amount to parallel open trials of two NF approaches, with some pre–post significant measures and a good many pre-FU significant measures for each approach. The most solid conclusion is that there was no deterioration for 6

months after Tx ended, but this must be tempered by uncertainty regarding intervening Tx. Finally, Leins et al.'s (2007) significant results are questionable as they were based on post hoc paired sample *t* tests performed after finding no significant Group × Time interactions, which is not standard statistical practice.

10. **McGrady et al. (2007)** presented results from a randomized study at the AAPB annual meeting. 31 American 7- to 12-year-olds (77% males, 32% White) with *T*-scores ≥65 on the ADHD Index of the CPRS-R (Conners, 2002a), comorbid with "depression, oppositional defiant, bipolar, intermittent explosive and obsessive–compulsive disorders" (p. 305). The presence of or change in concomitant Txs during the study was not reported. Participants were randomly assigned to either WLC or 30 theta/beta/SMR (unipolar Cz) NF sessions, for 45 to 60 min, 2 to 3 times per week over 10 to 15 weeks. No blinding of condition assignment was reported. Pre- and post-Tx assessments were given 2 weeks before and after Tx with a FU at the end of the school year. Those receiving NF, but not those on wait-list, improved significantly pre–post Tx on the CTRS-R (Conners, 2002b) ADHD ($p < .024$; $d = 0.40$) and Hyperactivity ($p < .011$; $d = 0.36$) subscales, and on the Gordon Diagnostic System (GDS; Gordon & Mettelman, 1988) computer-based Distractibility task ($p < .019$; $d = 0.81$). No significant differences between groups were found on pre–post Tx EEG theta and theta/beta scores, CTRS-R Inattention or Oppositional Behavior Scales, and the GDS Vigilance task or behavioral grades. Pre–post Tx improvement on the CTRS Hyperactivity Scales was reported to be maintained until the end of the school year but no statistics were reported. Finally, post-Tx EEG theta scores were significantly correlated with post-Tx and FU CTRS-R Hyperactivity (0.41, $p < .05$ and 0.43, $p < .05$, respectively) and post-Tx CTRS-R ADHD (0.41, $p < .05$) and GDS Distractibility (0.38, $p < .05$).

CRITIQUE: This was the first study in the field to include a majority (68%) non-White sample. Other study strengths include randomization, significant pre–post Tx changes on teacher-rating scales, and a computer-based test of attention. Unfortunately, there was no statistical analysis reported for FU, there was no blinding or sham-NF condition, the sample was not reliably diagnosed with ADHD, change in concomitant Txs during the study was not reported, and pre- and post-Tx assessments were not given immediately before and after Tx, which are all potential threats to the internal validity of the reported results.

362   *Journal of Attention Disorders 16(5)*

11. **Gevensleben, et al. (2009a)** published the first multisite randomized study of NF. They examined 94 German 9- to 12-year-olds (80% male) with *DSM-IV* ADHD (70% combined, 30% inattentive type; comorbidity: 23% dyslexia, 7% conduct disorder, 6% emotional disorder, 3% tics). All participants were medication free and without psychotherapy for at least 6 weeks before training (presence of other concomitant interventions and change in concomitant Tx during the study were not reported). Children were randomly assigned to one of two groups, NF ($n = 59$) or computerized attention skills training ("Skillies program," $n = 35$), using a 3:2 ratio. NF consisted of one block of theta/beta and one block of SCP training in a balanced order (unipolar Cz). Each block of NF involved eighteen 50-min, double sessions, 2 to 3 times per week, over 3 to 4 weeks. The design was semiblind, as teachers were blind but children and trainers were not. However, Gevensleben et al. noted that, although parents were not told of their child's Tx condition, only 42% and 37% of them in the NF and control groups, respectively, could not reliably identify Tx assignment. But this also means that 58% and 63%, respectively, did reliably identify Tx assignment indicating that the majority of parents were not blind.

The following ESs were reported by the authors. Compared with the control group, the NF group improved significantly more pre- to post-Tx for parent- and teacher-rated *DSM-IV* total ADHD ($p < .005$, $d = 0.60$; $p < .01$, $d = 0.64$, respectively) and inattentive symptoms ($p < .005$, $d = 0.57$; $p < .05$, $d = 0.50$, respectively) and for parent-rated hyperactive/impulsive ($p < .05$, $d = 0.45$), oppositional ($p < .05$, $d = 0.38$), and delinquent/physically aggressive ($p < .05$, $d = 0.37$) symptoms. On the parent- and teacher-rated Strengths and Difficulties Questionnaire (Woerner, Becker, & Rothenberger, 2004), the NF group had significantly greater decreases than the control group on hyperactivity ($p < .005$, $d = 0.60$; $p < .05$, $d = 0.48$, respectively) and on parent-rated overall scores ($p < .01$, $d = 0.51$). The responder rate (i.e., improvement >25% on parent-rated *DSM-IV* ADHD scale) in the NF group was significantly ($p < .05$) superior to the control group (51.7% vs. 28.6%). Regarding any differences between the two NF protocols, both theta/beta and SCP had comparable improvements on the parent-rated *DSM-IV* ADHD scale. In terms of parental evaluation of the NF and control Txs, there was no significant difference between groups in their attitude toward Tx effectiveness ($p = .77$) and their ratings of their child's motivation ($p = .71$).

That same year, Gevensleben, et al. (2009b) also reported distinct EEG effects from this study with combined theta/ beta and SCP NF training producing significant ($p < .001$) reductions in theta activity in centro-parietal regions compared with control. More specifically, for theta/beta NF decreases of theta activity over parietal-midline sites was significantly associated ($R = .465$, $p = .016$) with decreases in parent-rated *DSM-IV* ADHD total scores accounting for 20% of the variance. For SCP, NF increases of alpha over central-midline area was significantly associated ($R = 0.644$, $p < .001$) with decreases in parent-rated *DSM-IV* hyperactive/impulsive scores accounting for 40% of the variance. One year later, Gevensleben et al. (2010), published 6-month FU data on 61 (65%) of the participants who had not dropped out or been excluded because they had begun some other Tx. Overall, the NF group continued their pre–post Tx improvements over the control group on parent-rated (other measures not reported) *DSM-IV* ADHD total scores ($p < .005$, $d = 0.71$), inattentive scores ($p < .05$, $d = 0.73$), hyperactive/impulsive scores ($p < .01$, $d = 0.35$), and delinquent/aggressive scores ($p < .01$, $d = 0.52$), and on Strengths and Difficulties Questionnaire hyperactive scores ($p < .005$, $d = 0.49$). The NF group also improved on its pre–post Tx improvements at FU for homework problems ($p < .005$, $d = 0.60$). Finally, as only 50% of children were classified as responders (i.e., 25% reduction in parent-rated *DSM-IV* ADHD scores) and 19% of the NF group started medication during FU, Gevensleben et al. (2010) concluded that NF does not work for all children with ADHD and recommended it as one module in a multimodal Tx program rather that as a stand-alone Tx.

CRITIQUE: This study was significant in being the first to find distinct EEG effects related to two different types of NF, the first multisite study, and the largest randomized sample to date ($n = 94$). Other study strengths included *DSM* diagnoses, control of concomitant medication and psychotherapy, measurement of comorbidity, multidomain assessment, standard Tx outcome measures, EEG Tx outcome, examination of parental expectations and knowledge of Tx assignment, semiblind parents and teacher blindness, an examination of blinding validity, a comparison Tx, and FU. In fact, the use of a control group to teach attention skills may even have made it more challenging to obtain comparative significant results for NF group. Nevertheless, this study found significant results across multiple domains of assessments, including the core symptoms of ADHD at home and school, changes in the centro-parietal regions of the brain, and maintenance of parent-rated improvements at 6-month FU. Limitations included the confounding of both theta/beta training and SCP training in the same participants and lack of complete blinding. Yet, as we noted for Levesque et al. (2006), it seems unlikely (though possible) that expectancy biases would activate those particular brain areas in those receiving NF. Thus, this study, while an important advance, still leaves some uncertainty.

12. **Holtmann et al. (2009)** studied 34 German children, 7 to 12 years old (91% male) with an International Classification of Diseases–10th Revision (World Health Organization, 1993) hyperkinetic disorder disturbance of activity and attention (comparable with *DSM-IV* ADHD-combined type, 59%), hyperkinetic conduct disorder (comparable with *DSM-IV* ADHD-combined type with comorbid conduct disorder, 35%), and attention deficit disorder without hyperactivity (comparable with *DSM-IV* ADHD-inattentive type, 6%), diagnosed in a university psychiatric outpatient department. Except for conduct disorder, other comorbid diagnoses were not specifically assessed, although both Tx groups had CBCL scores in the borderline-clinical range for anxiety/depression, social, delinquent, and aggressive problems. All participants attended a 2-week school-based intensive behavioral day clinic, complimented with a weekly 1.5-hr parent-training session. Most (79%) children were on medication, kept constant during study, but no other information regarding the presence or change of concomitant Tx was reported.

Children were randomly assigned (3:2 ratio) to twenty 30-min sessions, twice a week, over 2 weeks, of theta/beta (unipolar, Cz) NF ($N = 20$) or to a cognitive-training control condition (Captain's Log program, $n = 14$). There was no attempt to blind. Both groups showed improvement on a Stop-Signal test, but only NF had a significant ($p = .001$) reduction of impulsivity errors, which was significantly better than control ($p = .018$, $d = 0.91$). On an event-related neurophysiologic measure of response inhibition (NoGo-N2), the NF group had a marginally significant ($p = .070$) increase in N2-amplitude (an indicator of NoGo-N2 normalization). Parent-rated SNAP-IV (Swanson, 1992) inattention, hyperactivity, and impulsivity showed improvements over time for both groups but no significant differential effects.

CRITIQUE: This study was important because it found that NF led to normalization of a key neurophysiologic correlate of response inhibition. Study strengths included randomization, *DSM*-compatible diagnoses, multidomain assessment, standard Tx outcome measures, measurement of comorbidity, and a control Tx. Like Gevensleben, et al.'s (2009a) study, the use of a control group to teach attention skills may have made it more challenging to obtain comparative significant results for the NF group. The challenge was further complicated by both groups receiving intensive all-day behavior therapy, which may have preempted some of the variance available to show a Tx group difference.

Unfortunately, this study lacked a triple-blind, sham control and FU. A final limitation is that only one significant result was reported for the core ADHD symptom of impulsivity.

13. **Urichuk et al. (2009)** presented on a randomized triple-blind study comparing real and sham NF at the Canadian Psychiatric Association annual meeting. Their sample included 37 (86% boys, ethnicity data unavailable) 7- to 15-year-old Canadians with *DSM-IV* ADHD-combined type (via *DSM* criteria on parent and teacher ADHD ratings, comorbidity data unavailable), 46% of whom were medicated (data on other concomitant Txs and change in all Txs during study not available). In all, 20 participants were randomized to a NF condition and 17 to a sham-NF condition (descriptions of NF and sham-NF Txs were unavailable). The study also included an education and a safety monitoring component for all participants, teachers, and families. Post-Tx blinding results indicated that most children (62%), teachers (62%), and NF trainers (100%) and nearly half of the parents (47%) were able to guess Tx condition. Children in both groups showed improved outcomes over the course of the project, but there were no significant (*ps* not available) differences between NF and sham NF. There were no significant changes in impulsivity or inattention levels of children on the TOVA for either group over the course of the study. Parents and children at each session reported no safety issues or adverse effects over the course of the study. Over time in both groups, parents reported children were doing better at the end, and children reported feeling more calm, being able to concentrate better, having better sleep quality, and feeling less discomfort in general. *M* and *SD* were unavailable, so ESs could not be calculated.

CRITIQUE: This study is significant in its use of a triple-blind and sham-NF condition, an examination of blinding validity, the lack of any significant differences between the NF and sham groups and the monitoring and reporting of adverse effects. Unfortunately, children, parents, raters, and NF trainers were not adequately blinded so their expectations could have affected the results. Additional limitations included lack of information on the presence of and change in concomitant Txs and the lack of FU data. Finally, like the preceding two randomized, triple-blind sham-NF studies (deBeus & Kaiser, 2010; Picard et al., 2006), this study has not undergone peer-reviewed publication.

14. The most recent randomized study of NF for pediatric ADHD was published by **Perreau-Linck et al. (2010)**. They examined nine 8- to 13-year-old Canadians (89% males, ethnicity not reported) with ADHD-combined type (no comorbidity), diagnosed by the Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Life Time Version (K-SADS-PL) semistructured interview (Kaufman et al., 1997), none of whom were on medication during the NF training (data on other concomitant Txs and change in all Txs during study not reported). Participants were randomized to an active-NF group ($n = 5$), receiving forty 60-min sessions of theta/SMR (unipolar C4 [central lobe, right hemisphere]) NF, 3 times a week for 79 weeks, or a sham-NF group ($n = 4$), receiving prerecorded sessions of the first author's EEG activity. Children, parents, and NF trainers were blind to group assignment but not the examiner conducting pre- and post-Tx neuropsychological testing. Participants in both groups showed significant (≥ 1.5 SD) individual pre–post Tx improvements on several CPRS-R (Conners, 2002a) subscales, particularly Hyperactivity, with more overall improvement in the sham group. All participants showed improvement on at least one of several neuropsychological measures, with more active-NF participants demonstrating improvement on the Stroop Task Inhibition/Switching Condition (Delis, Kaplan, & Kramer, 2001) and more sham-NF participants showing more improvement on the Stroop Task Inhibition Condition and the CPT-II Variability measure (Conners, 2002c). As noted by the authors, "The presence of placebo responses suggests that other factors, such as motivation or expectations, might contribute to the outcome of NF training in children with ADHD" (p. 230). M and SD were unavailable, so ESs could not be calculated.

CRITIQUE: This study is an important contribution to the field because of its use of a sham-NF group and triple blinding of children, parent raters, and NF trainers, and the lack of any significant differences between the NF and sham groups, Additional strengths include use of a reliable and valid diagnostic instrument, control of comorbidity and medication, and use of standard ADHD Tx outcome measures. However, the sample size was very small, there was no validity check of blinding, no FU data, and the results could have been affected by the assessment examiner not being blind and the presence of and change in concomitant Tx. In addition, the sham-NF condition may not have been completely inert and may have accidentally given active and effective feedback to participants.

Meta-Analysis of Published and Unpublished Trials

In 2009, Arns et al. published the first meta-analysis of 15 studies (6 from Germany and 5 from the United States, $N = 1,194$), 10 of which were prospective controlled studies ($n = 476$) and 5 pre–post design studies ($n = 718$). Among those 15, 4 were labeled as RCTs (Bakhshayesh, 2007; Gevensleben, Holl, Albrecht, Vogel, et al., 2009a; Holtmann et al., 2009; Levesque et al., 2006), and one other study included randomization to either theta/beta or SCP NF (Leins et al., 2007 & Strehl et al., 2006). In contrast to our review, Arns and colleagues included the study by Bakhshayesh (2007), an unpublished doctoral dissertation, but omitted the following 6 studies from their meta-analysis: Linden et al. (1996) because SD required for the meta-analysis were unavailable, and Fine et al. (1994), Palsson et al. (2001), Orlandi and Greco (2004), deBeus (2006), and Picard et al. (2006) because they were not published or part of a doctoral dissertation.

Pre- and post-Tx data used for the meta-analyses included rating scale information for inattention and hyperactivity and computerized tests (CPT, Go-No-Go test) for impulsivity. For the 10 controlled between-participant design studies, Arns and colleagues calculated mean ESs using Hedges D and fail-safe numbers (FSN), which are the number of unpublished null findings needed to render an effect nonsignificant to control for "file-draw" bias. They reported the following ESs and FSNs for inattention (0.81, 52.1), hyperactivity (0.40, 15.4), and impulsivity (0.69, 37.7). Within-participant ESs were also calculated for the controlled and within-participant (pre–post) design studies. For these, Arns et al. reported the following pre-post (uncontrolled) mean ESs and FSNs for inattention (1.02, 508.6), hyperactivity (0.71, 320.3), and impulsivity (0.94, 511.7).

Arns et al. (2009) also examined several additional post hoc questions including a comparison of NF and methylphenidate via three studies (none of which used randomization) and reported a mean ES for impulsivity of −0.034 with an FSN of 0 (there was not enough data to examine inattention or hyperactivity), suggesting to Arns and colleagues that NF may have similar effects to stimulant medication but because self-selection may have artificially inflated NF outcomes relative to medication, randomized comparisons (measuring all 3 ADHD symptom clusters) are needed before concluding relative effectiveness. (p. 5). For older (<2006) versus more recent studies (>2006), one-way ANOVA post hoc analyses of ESs did not reveal any significant differences nor were there any significant differences between the various NF protocols used. However, a significant correlation was reported for the average number of study sessions and improvement for attention ($p = .04$; $r = .550$) but not for impulsivity and hyperactivity, indicating more NF sessions lead to greater reductions in inattention. ESs for randomized ($n = 6$) and

nonrandomized ($n = 7$) studies were also compared. They found the ES for hyperactivity in randomized studies (0.54) was marginally significantly ($p = .080$) lower than for non-randomized studies (ES = 0.80) but no significant differences for inattention (randomized ES = 0.87, nonrandomized ES = 1.24) or impulsivity (randomized ES = 0.75, nonrandomized ES = 0.99; randomized and nonrandomized inattention and impulsivity ESs shared by M. Arns in a personal communication, June 29, 2010). These results suggested to Arns and colleagues that hyperactivity is more susceptible than inattention and impulsivity to nonspecific effects.

Arns et al. (2009) summarized their results by noting that "both prospective controlled studies and studies employing a pre- and post-design found large effects sizes (ES) for neurofeedback on impulsivity [$d = 0.69$] and inattention [$d = 0.81$] and a medium ES for hyperactivity [$d = 0.40$]" (p. 180). They concluded, "Neurofeedback treatment for ADHD can be considered 'Efficacious and Specific' (Level 5) with a large ES for inattention and impulsivity and a medium ES for hyperactivity" (p. 180).

CRITIQUE: This first meta-analysis of NF for pediatric ADHD was a timely and significant addition to the field. Arns and colleagues' approach to the meta-analyses, including their study search procedure, inclusion criteria, pre- and post-measure data identification, use of between-participant and within-participant ESs, FSNs, and post hoc analyses were all systematic, comprehensive, and informative. However, their use of both randomized and nonrandomized studies in their main meta-analyses is questionable because results from studies without randomization and control groups are difficult to interpret as they may be due to the specific effect of NF, and/or nonspecific Tx effects, selection effects, participant history, regression to the mean, placebo response, maturation, practice with assessment measures, and/or participant–rater–experimenter expectancies. Admirably, Arns and colleagues ran post hoc analyses to compare the ESs from randomized and nonrandomized studies but only one of the six randomized studies (Gevensleben, et al., 2009a) used data from partially blind raters (i.e., 37%-43% of parents who could not reliably identify Tx assignment). Without such blinding, participant–rater–experimenter expectancy effects may have contributed to the ESs reported by Arns et al. Therefore, due to the lack of controls, it is unclear as to whether the large ESs for impulsivity and inattention and the medium ES for hyperactivity are due to the active component of NF and/or nonspecific Tx factors.

We also question Arns et al.'s (2009) consideration of the semiactive control groups used in Bakhshayesh (2007), Gevensleben et al. (2009a), and Holtmann et al. (2009) as a "credible sham control providing an equal level of cognitive training and client-therapist interaction" (p. 180). A "sham" by definition entails a "fake" element imitating another condition such as an inactive substance (placebo) emulating a medication or a procedure/device

made to appear like a known Tx. Bakhshayesh used an electromyograph biofeedback control condition, whereas the other two studies used a computerized attention training control condition that did not involve placement of electrodes on participants' heads. Because all three studies used informed consent of the participants regarding the different procedures, it is doubtful any of the participants thought they were getting a "fake" form of NF. Hence, none of these procedures really can be considered as credible sham controls. However, these active-control conditions, which use interventions with assumed pre–post Tx gains, are theoretically a more stringent test of NF because they are able to subtract from the NF pre–post benefit whatever pre–post benefit the control Tx yielded, assuming that the control Txs are proven Txs. They should also theoretically control for a nonspecific placebo effect, provided the participants and investigators believe that they are as good as NF. However, it is doubtful that the investigational team believed they are as good; whether the parents, had similar effects to teachers, and other nonblind raters believed this depended heavily on the orientation and "sales pitch" given to them at the beginning and whether they believed this, the details of which are not presented. Furthermore, because the control apparatus and procedure are overtly different from those of NF, blinding was not possible. It is doubtful that any of the children, raters, or experimenters thought the control children were getting NF (i.e., "a credible placebo control") and so these studies could not have controlled for child-, rater-, and experimenter-expectancy effects.

Therefore, due to the inclusion of nonrandomized studies and the absence of studies with blinding and sham-NF designs, we disagree with Arns et al.'s (2009) conclusion that NF for pediatric ADHD be considered "efficacious and specific" (Level 5 of the AAPB, ISNR, and APA guidelines) as NF has NOT been "shown to be statistically superior to credible placebo therapies or to actual treatments . . . in two or more independent studies."

Summary of Study Designs and Findings

Research on NF for pediatric ADHD has evolved a great deal since the first randomized study was conducted in 1994 by Fine et al. This advance came at a rapid rate, with most (79%) of the studies conducted in a 5-year period between 2004 and 2010, impressive for any field, let alone one that traditionally has received little funding. All of the studies used theta/beta training, one used theta/beta or SCP, and another theta/beta and SCP. Where data were available (50% of studies), the overall mean ES for all Tx outcome measures was 0.69 (medium ES, range = 0.34-1.66); ADHD measures 0.69 (medium ES, range = 0.34-1.62), inattention measures 0.79 (medium ES, range = 0.41-1.62), and hyperactivity/impulsivity measures 0.71 (medium ES,

range = 0.35-1.55). Five studies also showed neurophysiological changes specifically associated with NF Tx (EEG; deBeus & Kaiser, 2011, Gevensleben, et al., 2009b, Palsson et al., 2001; fMRI; Levesque et al., 2006; and N2-amplitude: Holtmann et al., 2009). Regarding control groups, six (43%) used wait-list, four (29%) sham NF, four (29%) non-NF cognitive training, and two (14%) an alternative NF procedure. A total of 8 (57%) studies used a blinding design for informants (i.e., parent-, teacher-, clinician-rater), 5 (36%) for children, 4 (29%) for NF trainers, and 6 (43%) did not use any blinding. Only four (29%) used a triple blind (i.e., children, informants, and trainers) and two (14%) a double blind.

On average, NF was given for approximately 46 min (range = 30-60 min), two and half times per week (range = 1-5/week), over 32 sessions (range = 18-40), during the course of 13 weeks (range = 2-24 weeks). The majority (64%) of studies used a unipolar electrode, with most (67%) of those adopting a Cz placement. Various programs were used but only two (SmartBrain and Thought Technology) were used more than twice (three studies each). Six (43%) studies have been conducted in the United States, five (36%) in Canada, and three (21%) in Germany. In total, 504 participants ($M = 36$, range = 9-94) have been examined, between the ages of 5 and 15, mostly (64%) with children ≤ 12 years of age, males (84%), and, where ethnicity was reported (36% of studies), Whites (82%). Most (79%) studies used an evidence-based assessment to diagnose ADHD. All but three studies (79%) used *DSM-IV* criteria to diagnose ADHD, with most (79%) participants having the combined type (inattentive type = 11% and hyperactive/impulsive type = 10% of studies) and, where reported (50% of studies), with various comorbidities (M comorbidity per study = 16% of sample). Regarding concomitant Txs, medications were reported for most (86%) of the studies, with a M of 41% of participants on medication in those studies, but only three studies (21%) reported other concomitant Txs such as psychotherapy and special education. Only four studies (29%) tracked changes in concomitant medication during the study and only two (14%) tracked changes in other concomitant Txs. Regarding multidomain assessment of Tx outcome (e.g., parent-, teacher-, or clinician-ratings, neuropsychological testing, neurophysiological measures), 50% of studies used measures in two domains and 50% in three/four domains, but all studies used at least one measure (or a German equivalent) typically used in Tx outcome research on pediatric ADHD.

Therefore, overall study strengths include randomization (100% of studies), use of evidenced-based assessments to diagnose ADHD (79%), *DSM* criteria (79%), standard Tx outcome measures (100%), multidomain assessment (100% with ≥ 2 domains), identification of medication as a concomitant Tx (79%), moderate-sized samples ($M = 36$), non-WLC condition (71%), and some type of blind (57%).

Summary of Study Limitations

1. Triple Blinding:

Just over half (57%) of the studies adopted some type of blinding design, with only four (29%) of these a triple-blind procedure to attempt to control for participant, rater, and NF-trainer expectancies of Tx outcome (deBeus, 2006/deBeus & Kaiser, 2011; Perreau-Linck et al., 2010; Picard et al., 2006; Urichuk et al., 2009). Admittedly, triple-blinding participants, raters, and Tx providers is comparatively easier with medication where an inert placebo pill can be used. However, even in medication studies, the validity of blinding has been questioned and requires actual testing of the accuracy of blinding during the early, mid, and later stages of Tx (Margraf et al., 1991). In contrast, in psychotherapy Tx outcome studies, where knowledge of the Tx is required for a therapist to administer it, a placebo condition is virtually impossible. In the NF field, one practical reason for not using a triple-blind design earlier was that, with the NF technology available, the trainer had to know which parameters to train to adjust the feedback. It appears that the required technology to conduct NF without a trainer was not available to the field until 2006. In that year, deBeus and Picard et al. used NF training devices with interface modules programmed to give feedback contingent on the participant's EEG, which eliminated the need for a trainer to monitor and adjust the feedback parameters, and therefore allowed the experimenter/trainer to be blind. Although this technology has been available since 2006, since then only two more studies have used a triple-blind design. However, only two of the studies (Gevensleben, et al., 2009a; Urichuk et al., 2009) tested the accuracy of their blinding strategy as recommended for medication studies by Margraf et al. (1991). Unfortunately, both studies found their blinding was suboptimal as most parents were able to identify Tx assignment in Gevensleben, et al.'s study (NF = 58%, control group = 63%) and Urichuk et al.'s (2009) study (children = 62%, teachers = 62%, NF trainers = 100%, and parents = 47%).

2. Sham-NF Condition:

Another reason for the lack of triple-blind studies is probably because only two more studies (Perreau-Linck et al., 2010; Urichuk et al., 2009) since 2006 have used a "true" sham-NF control condition. However, as only Urichuk et al. (2009) tested their blind but found it inadequate, no studies have yet used a valid sham-NF condition and valid triple-blind design and therefore adequately controlled for all nonspecific Tx effects and isolated a relatively untainted specific measure of the active component of NF. Therefore, the overall ADHD mean ES we previously calculated (0.69) may well be a measure of the size of the NF effect or the size of the effect of nonspecific factors or a combination of both or the difference between a less potent Tx and NF.

Lofthouse et al. 367

Apart from the absence of the required technology for the inclusion of a sham-NF condition, another reason for this methodology not being included in Tx outcome research in this area may be because of ethical concerns. Based on the ethical principles outlined in the Declaration of Helsinki (World Medical Association, 2000), La Vaque and Rossiter (2001) were one of the first to argue against sham- or placebo-controlled studies in NF research on pediatric ADHD. They noted that the Helsinki principles "specifically prohibit designs that would withhold or deny 'the best proven diagnostic and therapeutic' Tx to any participant in a clinical study, including those individuals who consent to randomization into a control group" (p. 23). La Vaque and Rossiter instead suggested active-control studies with NF compared with an intervention with known clinical efficacy (e.g., stimulant medication, atomoxetine, behavior modification). However, others have argued against this recommendation for several reasons including the Declaration of Helsinki not being "accepted as the world ethical standard" (Striefel, 2001, p. 39); those principles being "aspirational rather than mandatory" (Striefel, 2001, p. 39); use of placebo studies when there is minimal risk, the scientific necessity of a placebo when care is taken to ensure the valid informed consent, and protection of the rights and welfare of participants as specified in other relevant ethical codes, principles, or laws and institutional review boards (Saunders & Wainwright, 2003; Striefel, 2001; Vrhovac, 2004); the difficulty of defining and identifying "best-proven" Txs (Glaros, 2001); the prevention of important research and future clinical care (Saunders & Wainwright, 2003; Vrhovac, 2004); and the possibility that placebo-controlled designs may actually require fewer participants than active controls and therefore reduce exposure to the risks of nonresponse and adverse effects (Young, 2001). In addition, the objections to using a sham Tx seem to assume that NF is a "best-proven Tx," which it is not, and overlook the fact that many families refuse stimulant medication, which is a "best-proven Tx." On balance, the ethical arguments against a sham placebo condition are not compelling.

Some ask why research on NF should require a sham condition (or a triple-blinded Tx control) and thus be held to a higher standard than psychotherapy outcome research. We believe that any Tx offered to the public that does not have any practical, ethical, or risk-related obstacles to prevent the use of a sham condition and triple-blinding design should be scientifically examined at the highest level. Another reason the efficacy of NF should be examined in this manner is because of its expense and laborious intensity. Therefore, we believe that, when there is minimal risk and steps are taken to ensure and monitor the valid informed consent and protection of participants' rights and welfare, it may be actually unethical to fail to test, at the highest level of scientific scrutiny, an expensive, time-consuming,

and hope-inspiring Tx that is being marketed to the public as effective.

Finally, like with blinding, the validity of any sham condition should be examined and reported. Although the ideal sham-NF condition is setup to give random feedback not contingent on the child's EEG, it might become inadvertently associated with the child's EEG and thus give actual feedback. This could happen in two ways. First, sham-NF participants will, at various times during their sessions, actually be in the "correct" EEG bandwidth when random feedback occurs, thus reinforcing that behavior. Second, because all children in both the sham and real-NF groups are told certain feedback means you are not paying attention and need to, random feedback could cue children into paying more attention thus reinforcing this behavior. Currently there are no data in the field to test these hypotheses but future research needs to examine these possibilities.

3. Identification, Measurement, and Control of Concomitant Txs:

Although most (86%) studies identified the presence of concomitant medication, only three studies identified concomitant psychotherapy and/or educational Txs. However, only four studies (29%) tracked changes in concomitant medication during the study and only two studies (14%) monitored changes in other concomitant Txs. Therefore, resultant positive changes associated with NF could be caused, moderated, or even mediated by unreported changes in concomitant Tx.

4. Post-Tx FU:

FU studies in any Tx study are crucial for examining long-term efficacy and/or the need for frequency and duration of booster sessions. To date, only three studies (Gevensleben et al., 2010; Leins et al., 2007; McGrady et al., 2007) used a post-Tx FU assessment of NF. Unfortunately, although McGrady et al. reported pre–post maintenance of CTRS hyperactivity until the end of the school year, no statistics were reported. Although both Leins et al. and Gevensleben et al. reported positive results at 6 months FU, they did not identify, control, or measure potential confounds during this FU period. In addition, as only 4/17 significant results in Leins et al. became significant between baseline and immediately post-Tx (the remaining 13 becoming significant between baseline and 6-month FU), unless there was a delayed "sleeper effect" of NF, the positive FU results are more likely due to events after the NF Tx was terminated. The Gevensleben et al. study was also limited by not being fully blinded, not using a sham NF, not identifying/tracking concomitant Txs, and only reporting parent ratings at FU.

5. Identification, Monitoring, and Reporting of Adverse Side Effects:

Although NF is frequently reported as safe and without side effects, only Urichuk et al. (2009) measured and reported adverse events noted by children and parents related to NF, and they found no such adverse effects. This element of Tx outcome research is essential because, as Loo and Barkley (2005) noted, "All truly effective treatments produce some side effects in some percentage of the population . . . because individuals differ in their physiological makeup, particularly brain organization and functioning" and because "clinical ineptitude in the delivery of the treatment in some cases and as a consequence of comorbid disorders in other cases always ensures that some patients will not respond well to the intervention as delivered" (p. 74). Furthermore, none of the reviewed studies identified or monitored any adverse events related to the control condition. This is just as important as Birch (2006) demonstrated in research on surgery and acupuncture, sham conditions may actually not be inert and so called for researchers to "stop stating they are using placebo-controlled procedures unless they can present evidence that their sham is in fact inert and valid as a placebo treatment" (p. 307). Therefore, similar to Margraf et al.'s (1991) recommendations for testing the accuracy of blinding, the validity of sham conditions should also be evaluated during the early, mid, and later stages of Tx.

6. Large Variability in the Delivery of NF:

A final limitation of existing studies concerns the lack of a standard method for giving NF to children with ADHD. In the 14 reviewed studies, there was a wide variety in the technology used (only two devices were used more than once), electrode placements (50% Cz, 50% other locations), session number (18-40), duration (30-60 min), frequency (1-5 times a week), and course (2-24 weeks).

Future Research Directions

Based on our review and identification of limitations of the research on the NF Tx of pediatric ADHD, the primary and essential area for improvement involves peer-reviewed and published randomized, triple-blind, sham/Tx-controlled trials. As previously noted, a sham-NF condition or a triple-blinded Tx-control condition requires use of technology or a design that eliminates the need for a trainer to adjust the feedback/Tx parameters to keep all involved blind. We are not saying that all NF research should involve this type of design without the trainer, but that it is an imperative step, along with randomization and triple blinding, to demonstrate the efficacy of the primary active component of feedback. If NF is shown to be *"efficacious and specific,"* then

additional Tx components, such as a trainer adjusting the feedback, can be systematically examined and compared with the original efficacy studies to evaluate the effects of additional Tx components above and beyond the primary feedback component.

Due to the previously discussed limitations, these initial efficacy trials will also require testing of the accuracy of blinding and sham condition and the continuous monitoring of potential adverse events related to the NF and sham conditions. Concomitant Txs, such as psychiatric medication, psychotherapy, and educational services that could confound study results also need to be identified, monitored, and controlled (methodologically or statistically) during the study. For any Tx, but particularly one that is as expensive and intensive as NF, long-term gains, above and beyond post-Tx, need to be shown, along with the identification, control, or measurement of potential confounds during the FU period.

If NF research achieves Level 5 status ("efficacious and specific") for a specifically defined population, researchers can then investigate its efficacy for a variety of other groups and conditions seldom included in the initial efficacy trials. For example, investigators could examine NF Tx for ADHD for non-Americans/Canadians, females, adolescents, non-Whites, *DSM-IV* ADHD-inattentive and hyperactive/impulsive types, and participants receiving concomitant Tx. Level 5 status also offers researchers a stronger position from which to conduct comparison studies with medication to see whether NF has similar or greater effects than conventional Tx (i.e., an alternative Tx) or incremental effects when added to conventional Tx (i.e., a complimentary Tx). Additional studies can also examine the necessary number, length, frequency per week, and overall Tx duration of NF sessions required to obtain clinical improvement and to sustain that improvement over time. A final issue for resolution by future research is whether all types of NF are equally effective for pediatric ADHD or just those approaches found to be efficacious in randomized, triple-blind sham-controlled trials. Therefore, future research on NF for pediatric ADHD needs to identify which specific NF approaches (e.g., theta/beta, SMR, SCP), electrode number and placement (e.g., one vs. multiple channels, Cz vs. other placement), and specific NF program works and for which participants.

Conclusion

Although the evidence for the NF Tx of pediatric ADHD is increasing in quantity and quality, because of some questions about methodology in most studies, the evidence is promising but not yet conclusive. Due to the lack of blinding and sham-control conditions in randomized studies, we disagree with Arns et al. (2009) who concluded that "neurofeedback treatment for ADHD can be considered 'Efficacious and Specific' (level 5)" (p. 188). Instead,

based on the guidelines used by Arns et al. (La Vaque et al., 2002) and APA (Chambless et al., 1998) and our research review, we conclude that NF Tx for pediatric ADHD can be currently considered as "probably efficacious" and in need of a large multisite triple-blind sham-controlled RCT to settle the issue.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

References

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR, & TRF profiles*. Burlington: University of Vermont.

Altepeter, T. S., & Breen, M. J. (1989). The Home Situations Questionnaire (HSQ) and the School Situations Questionnaire (SSQ): Normative data and an evaluation of psychometric properties. *Journal of Psychoeducational Assessment, 7*, 312-322.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (DSM-III, rev.)*. Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (DSM-IV, text rev.)*. Washington, DC: Author.

Arnold, L. E. (2004). *Contemporary diagnosis and management of ADHD* (3rd ed.). Newtown, PA: Handbooks in Health Care Co.

Arns, M., de Ridder, S., Strehl, U., Breteler, M., & Coenen, A. (2009). Efficacy of neurofeedback treatment in ADHD: The effects on inattention, impulsivity and hyperactivity: A meta-analysis. *EEG & Clinical Neuroscience, 40*, 180-189.

Bakhshayesh, A. R. (2007). *The efficacy of neurofeedback compared to EMG biofeedback in the treatment of ADHD children* (Unpublished doctoral dissertation). University Potsdam, Germany.

Banaschewski, T., Brandeis, D., Heinrich, H., Albrecht, B., Brunner, E., & Rothenberger, A. (2004). Questioning inhibitory control

as the specific deficit of ADHD-evidence from brain electrical activity. *Journal of Neural Transmission, 111*, 841-864.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of reporting source and definition of disorder. *Journal of Abnormal Psychology, 111*, 279-289.

Barry, R. J., Clarke, A. R., & Johnstone, S. J. (2003). A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. *Clinical Neurophysiology, 114*, 171-183.

Beauregard, M. & Levesque, J. (2006). Functional magnetic resonance imaging investigation of the effects of neurofeedback training on the neural bases of selective attention and response inhibition in children with attention-deficit/hyperactivity disorder. *Applied Psychophysiology and Biofeedback, 31*(1), 3-20.

Berger, H. (1929). About the human electroencephalogram. *Archives of Psychiatry and Nervous Diseases, 87*, 527-570.

Birch, S. (2006). A review and analysis of placebo treatments, placebo effects, and placebo controls in trials of medical procedures when sham is not inert. *Journal of Alternative and Complementary Medicine, 12*, 303-310.

Bush, G., Valera, E. M., & Seidman, L. J. (2005). Functional neuroimaging of attention-deficit/hyperactivity disorder: A review and suggested future directions. *Biological Psychiatry, 57*, 1273-1284.

Bush, P. J., Whalen, B. R., Rosen, M. A., Jenike, S. C., McInerney, S. L., & Rauch, S. L. (1998). The counting Stroop: An interference task specialized for functional neuroimaging–validation study with functional MRI. *Human Brain Mapping, 6*, 270-282.

Centers for Disease Control and Prevention. (2005). Prevalence of diagnosed and medicated attention-deficit/hyperactivity disorder: United States, 2003. *Morbidity & Mortality Weekly Report, 54*, 842-847.

Chabot, R. A., Merkin, H., Wood, L. M., Davenport, T. L., & Serfontein, G. (1996). Sensitivity and specificity of QEEG in children with attention deficit or specific developmental learning disorders. *Clinical Electroencephalography, 27*, 26-34.

Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., & Woody, S. R. (1998). An update on empirically validated therapies, II. *Clinical Psychologist, 51*, 3-16.

Clarke, A. R., Barry, R. J., McCarthy, R., & Selikowitz, M. (2001). Electroencephalogram differences in two subtypes of attention-deficit/hyperactivity disorder. *Psychophysiology, 38*, 212-221.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Conners, C. K. (2002a). *Conners' Parent Rating Scale–Revised (CPRS-R)*. Toronto, Ontario, Canada: Multi-Health Systems.

Conners, C. K. (2002b). *Conners' Teacher Rating Scale–Revised (CTRS-R)*. Toronto, Ontario, Canada: Multi-Health Systems.

Conners, C. K. (2002c). *CPT-II, Continuous Performance Test II.* Toronto, Ontario, Canada: Multi-Health Systems.

deBeus, R. (2006, May). *Progress in efficacy studies of EEG biofeedback for ADHD.* Paper presented at the meeting of the American Psychiatric Association, Toronto, Canada.

deBeus, R., & Kaiser, D. A. (2011). Neurofeedback with children with attention-deficit hyperactivity disorder: A randomized double-blind placebo-controlled study. In R. Coben & J. R. Evans (Eds.), *Neurofeedback and neuromodulation techniques and applications* (pp. 127-152). London, England: Academic Press.

Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis-Kaplan executive function system examiner's manual.* San Antonio, TX: Harcourt Assessment.

Donovan, H. S., Kwekkeboom, K. L., Rosenzweig, M. Q., & Ward, S. E. (2009). Nonspecific effects in psychoeducational intervention research. *Western Journal of Nursing Research, 31,* 983-998.

Drechsler, R., Straub, M., Doehnert, M., Heinrich, H., Steinhausen, H., & Brandeis, D. C. (2007). Controlled evaluation of a neurofeedback training of slow cortical potentials in children with attention deficit/hyperactivity disorder. *Behavioral and Brain Functions, 3,* 35.

DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale-IV: Checklists, norms, and clinical interpretation.* New York, NY: Guilford.

Durup, G., & Fessard, A. (1935). I. L'e´lectrence´-phalogramme de l'homme. Observations psycho-physiologiques relatives a` l'action des stimuli visuels et auditifs. *L'anne´e Psychologique, 36*(1), 1-32.

Electroencephalography. (2010, August 13). *In Wikipedia, The Free Encyclopedia.* Retrieved from http://en.wikipedia.org/w/index.php?title=Electroencephalography&;oldid=378780428

Eyberg, S. M., & Pincus, D. (1999). *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory–Revised.* Odessa, FL: Psychological Assessment Resources.

Faraone, S. V., Perlis, R. H., Doyle, A. E., Smoller, J. W., Goralnick, J., Holmgren, M. A., & Sklar, P. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57,* 1313-1323.

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology, Research and Practice, 40,* 532-538.

Fine, A. H., Goldman, L., & Sandford, J. (1994, August). *Innovative techniques in the treatment of ADHD: An analysis of the impact of EEG biofeedback training and a cognitive computer generated training.* Paper presented at the meeting of the American Psychological Association, Los Angeles, CA.

Fuchs, T., Birbaumer, N., Lutzenberger, W., Gruzelier, J. H., & Kaiser, J. (2003). Neurofeedback treatment for attention-deficit/hyperactivity disorder in children: A comparison with methylphenidate. *Applied Psychophysiology and Biofeedback, 28,* 1-12.

Gevensleben, H., Holl, B., Albrecht, B., Vogel, C., Schlamp, D., Kratz, O., & Heinrich, H. (2009a). Is neurofeedback an efficacious treatment for ADHD? A randomized controlled clinical trial. *Journal of Child Psychiatry and Psychology, 50,* 780-789.

Gevensleben, H., Holl, B., Albrecht, B., Schlamp, D., Kratz, O., Studer, P., . . . Heinrich H. (2009b). Distinct EEG effects related to neurofeedback training in children with ADHD: A randomized controlled trial. *International Journal of Psychophysiology, 74,* 149-157.

Gevensleben, H., Holl, B., Albrecht, B., Schlamp, D., Kratz, O., Studer, P., & Heinrich H. (2010). Neurofeedback training in children with ADHD: 6-Month follow-up of a randomised controlled trial. *European Child & Adolescent Psychiatry, 19,* 715-724.

Glaros, A. G. (2001). A comment on La Vaque and Rossiter. *Applied Psychophysiology and Biofeedback, 26,* 61-65.

Gordon, M., & Mettelman, B. (1988). The assessment of attention: I. Standardization and reliability of a behavior-based measure. *Journal of Clinical Psychology, 44,* 682-690.

Goyette, C. H., Conners, C. K., & Ulrich, R. F. (1978). Normative data on revised Conners parent and teacher rating scales. *Journal of Abnormal Child Psychology, 6,* 221-236.

Guy, W. (1976). *ECDEU Assessment Manual for Psychopharmacology—Revised* (DHEW Publ No ADM 76 338). Rockville, MD: U.S. Department of Health, Education, & Welfare, Public Health Service, Alcohol, Drug Abuse, & Mental Health Administration, NIMH Psychopharmacology Research Branch, Division of Extramural Research Programs.

Heinrich, H., Gevensleben, H., Freisleder, F. J., Moll, G. H., & Rothenberger, A. (2004). Training of slow cortical potentials in attention-deficit/hyperactivity disorder: Evidence for positive behavioral and neurophysiological effects. *Biological Psychiatry, 55,* 772-775.

Hirshberg, L. M., Chiu, S., & Frazier, J. A. (2005). Emerging brain-based interventions for children and adolescents: Overview and clinical perspective. *Child and Adolescent Psychiatric Clinics of North America, 14,* 1-19.

Holtmann, M., Grasmann, D., Cionek-Szpak, E., Hager, V., Panzer, N., Beyer, A., & Stadler, C. (2009). Specific effects of neurofeedback on impulsivity in ADHD. *Kindheit und Entwicklung, 18,* 95-104.

Kamiya, J. (1968). Conscious control of brain waves. *Psychology Today, 1,* 57-60.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test manual.* Circle Pines, MN: American Guidance Service.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., . . . Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Life Time Version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry, 36,* 980-988.

La Vaque, T. J., Hammond, D. C., Trudeau, D., Monastra, V. J., Perry, J., & Lehrer, P. (2002). Template for developing guidelines for the evaluation of the clinical efficacy of psychophysiological interventions. *Applied Psychophysiology and Biofeedback, 27,* 273-281.

La Vaque, T. J., & Rossiter T. (2001). The ethical use of placebo controls in clinical research: The Declaration of Helsinki. *Applied Psychophysiology and Biofeedback, 26,* 25-39.

Leark, R. A., Greenberg, L. K., Kindschi, C. L., Dupuy, T. R., & Hughes, S. J. (2007). *Test of variables of attention: Clinical manual.* Los Alamitos, CA: TOVA Company.

Leins, U., Goth, G., Hinterberger, T., Klinger, C., Rumpf, N., & Strehl, U. (2007). Neurofeedback for children with ADHD: A comparison of SCP and theta/beta protocols. *Applied Psychophysiology and Biofeedback, 32,* 73-88.

Levesque, J., Beauregard, M., & Mensour, B. (2006). Effect of neurofeedback training on the neural substrates of selective attention in children with attention-deficit/hyperactivity disorder: A functional magnetic resonance imaging study. *Neuroscience Letters, 394,* 216-221.

Linden, M., Habib, T., & Radojevic, V. (1996). A controlled study of the effects of EEG biofeedback on cognition and behavior of children with attention deficit disorder and learning disabilities. *Biofeedback & Self-Regulation, 21,* 35-49.

Loomis, A. L., Harvey, E. N., & Hobart, G. (1936). Electrical potentials of the human brain. *Journal of Experimental Psychology, 19,* 249.

Loo, S. K., & Barkley, R. A. (2005). Clinical utility of EEG in attention deficit hyperactivity disorder. *Applied Neuropsychology, 12,* 64-76.

Lubar, J. F., & Shouse, M. N. (1976). EEG & behavioral changes in a hyperkinetic child concurrent with training of the sensorimotor rhythm (SMR): A preliminary report. *Biofeedback & Self-Regulation, 3,* 293-306.

Mann, C., Lubar, J., Zimmerman, A., Miller, C., & Muenchen, R. (1992). Quantitative analysis of EEG in boys with attention-deficit-hyperactivity disorder: Controlled study with clinical implications. *Pediatric Neurology, 8,* 30-36.

Margraf, J., Ehlers, A., Roth, W. T., Clark, D. B., Sheikh, J., Agras, W. S., & Taylor, C. B. (1991). How "blind" are double-blind studies? *Journal of Consulting and Clinical Psychology, 59,* 184-187.

McGrady, A., Prodente, C., Fine, T., & Donlin, J. (2007, February). *Neurofeedback in children with behavioral and emotional problems.* Paper presented at the annual meeting of the Association for Applied Psychophysiology & Biofeedback. Monterey, CA.

Michelson, D., Allen, A. J., Busner, J., Cascat, C., Dunn, D., Kratochvil, C., & Harder, D. (2002). Once-daily atomoxetine treatment for children and adolescents with attention deficit hyperactivity disorder: A randomized, placebo-controlled study. *American Journal of Psychiatry, 159,* 1896-1901.

Monastra, V. J. (2005). Electroencephalographic biofeedback (neurotherapy) as a treatment for ADHD: Rationale and empirical foundation. *Child and Adolescent Psychiatric Clinics of North America, 14,* 55-82.

Monastra, V. J., Lubar, J. F., Linden, M., VanDeusen, P., Green, G., Wing, W., . . . Fenger, T. N. (1999). Assessing attention-deficit hyperactivity disorder via quantitative electroencephalography: An initial validation study. *Neuropsychology, 13,* 424-433.

Monastra, V. J., Monastra, D. M., & George, S. (2002). The effects of stimulant therapy, EEG biofeedback, and parenting style on the primary symptoms of attention-deficit/hyperactivity disorder. *Applied Psychophysiology and Biofeedback, 27,* 231-249.

MTA Cooperative Group. (1999). A 14-Month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry, 56,* 1073-1086.

Nall, A. (1973). Alpha training and the hyperactive child: Is it effective? *Academic Therapy, 9,* 5-19.

Orlandi, M. A., & Greco, D. (2004, August). *A randomized, double-blind clinical trial of EEG neurofeedback treatment for attention-deficit/hyperactivity disorder.* Paper presented at the meeting of the International Society for Neuronal Regulation, Fort Lauderdale, FL.

Palsson, O. S., Pope, A. T., Ball, J. D., Turner, M. J., Nevin, S., & deBeus, R. (2001, March). *Neurofeedback videogame ADHD technology: Results of the first concept study.* Paper presented at the annual meeting of Association for Applied Psychophysiology & Biofeedback, Research Triangle Park, NC.

Perreau-Linck, E., Lessard, N., Levesque, J., & Beauregard, M. (2010). Effects of neurofeedback training on inhibitory capacities in ADHD children: A single blind randomized placebo-controlled study. *Journal of Neurotherapy, 14,* 229-242.

Picard, C., Moreau, G., Guay, M. C., & Achim, A. (2006, September). *Double double-blind sham study of neurofeedback treatment in children with ADHD.* Paper presented at the meeting of the International Society for Neurofeedback and Research, Atlanta, GA.

Pliszka, S. (2007). Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 46,* 894-921.

Reich, W., Leacock, N., & Shanfeld, K. (1997). *DICA-IV, Diagnostic Interview for Children and Adolescents-IV.* Toronto, Ontario, Canada: Multi-Health Systems.

Reynolds, C. R., & Kamphaus, R. W. (1992). *Behavior assessment system for children: Manual.* Circle Pines, MN: American Guidance.

Robbins, J. (2008). *A symphony in the brain: The evolution of brain wave biofeedback—Revised edition.* New York, NY: Grove Press.

Rockstroh, B., Elbert, T., Lutzenberger, W., & Birbaumer, N. (1990). Biofeedback: Evaluation and therapy in children with attentional dysfunctions. In A. Rothenberger (Ed.), *Brain and behavior in child psychiatry* (pp. 345-355). Berlin, Germany: Springer.

Rossiter, T. R., & La Vaque, T. J. (1995). A comparison of EEG biofeedback and psychostimulants in treating attention deficit/hyperactivity disorders. *Journal of Neurotherapy, 1,* 48-59.

Sandford, J. A., & Turner, A. (1995). *Manual for the integrated visual and auditory continuous performance test.* Richmond, VA: Braintrain.

Saunders, J., & Wainwright, P. (2003). Risk, Helsinki 2000 and the use of placebo in medical research. *Clinical Medicine, 3,* 435-439.

Sheslow, D., & Adams, W. (1990). *WRAML: Wide range assessment of memory and learning.* Wilmington, DE: Jastak.

Song, D. H., Shin, D. W., Jon, D. I., & Ha, E. H. (2005). Effects of methylphenidate on quantitative EEG of boys with attention-deficit hyperactivity disorder in continuous performance test. *Yonsei Medical Journal, 46,* 34-41.

Sowell, E. R., Thompson, P. M., Welcome, S. E., Henkenius, A. L., Toga, A. W., & Peterson, B. S. (2003). Cortical abnormalities in children and adolescents with attention-deficit hyperactivity disorder. *Lancet, 362,* 1699-1707.

Spetie, L., & Arnold, L. E. (2007). Attention deficit hyperactivity disorder. In A. Martin & F. R. Volkmar (Eds.), *Lewis' child and adolescent psychiatry: A comprehensive textbook* (4th ed., pp. 430-453). Baltimore, MD: Wolters Kluwer/Lippincott Williams & Wilkins.

Strehl, U., Leins, U., Goth, G., Klinger, C., Hinterberger, T., & Birhaumer, N. (2006). Self-regulation of slow cortical potentials: A new treatment for children with attention-deficit/hyperactivity disorder. *Pediatrics, 118,* 1530-1540.

Striefel, S. (2001). Ethical research issues: Going beyond the Declaration of Helsinki. *Applied Psychophysiology and Biofeedback, 26,* 41-61.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Swanson, J. M. (1992). *School based assessments and interventions for ADD Students.* Irvine, CA: KC.

Swanson, J. M., & Castellanos, F. X. (2002). Biological bases of ADHD: Neuroanatomy genetics, and pathophysiology. In P. S. Jensen & J. R. Cooper (Eds.), *Attention deficit hyperactivity disorder: State of the science best practices* (pp. 7-20). Kingston, NJ: Civic Research Institute.

Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B., . . . Wu, M. (2001). Clinical relevance of the primary findings of the MTA: Success rates based on severity of ADHD and ODD symptoms at the end of treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 40,* 168-179.

Swanson, J. M., Nolan, W., & Pelham, W. (1981). *The SNAP Rating Scale.* Resources in Education.

Tewes, U., Rossmann, P., & Schallberger, U. (1999). *Hamburg Wechsler Intelligence Test for Children–Third Edition (HAWIK III).* Bern, Switzerland: Huber.

Urichuk, L., Hnatko, G., Baydala, L., Wikman, E., Hoover, H. J., Vohra, S., & Schachter, H. (2009, September). *Neurofeedback treatment of ADHD—A feasibility study.* Paper presented at the meeting of the Canadian Psychiatric Association, St. John's, Canada.

Vrhovac, B. (2004). Placebo and the Helsinki declaration—What to do? *Science and Engineering Ethics, 10,* 81-93.

Walter, W. G., Cooper, R., Aldridge, V., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electrical sign of sensorimotor association and expectancy in the human brain. *Nature, 203,* 380-385.

Wechsler, D. (1974). *Examiner's manual: Wechsler Intelligence Scale for Children–Revised.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children (4th ed.): Technical and Interpretive Manual.* San Antonio, TX: Psychological Corporation.

Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychological journals: Guidelines and explanations. *American Psychologist, 54,* 594-604.

Woerner, W., Becker, A., & Rothenberger, A. (2004). Normative data and scale properties of the German parent SDQ. *European Child & Adolescent Psychiatry, 13,* 11-16.

Woodruff, T. J., Axelrad, D. A., Kyle, A. D., Nweke, O., Miller, G. G., & Hurley, B. J. (2004). Trends in environmentally related childhood illnesses. *Pediatrics, 113,* 1133-1140.

World Health Organization. (1993). *The ICD-10 classification of mental and behavioural disorders. Diagnostic criteria for research.* Geneva, Switzerland: Author.

World Medical Association. (2000). *The Declaration of Helsinki 52nd WMA General Assembly.* Edinburgh, Scotland: Author. Available from http://www.wma.net

Young, L. D. (2001). Editorial: The ethics of placebo controlled clinical research and applied psychophysiology. *Applied Psychophysiology and Biofeedback, 26,* 17-21.

Bios

Nicholas Lofthouse is an assistant professor–clinical in the Department of Psychiatry at The Ohio State University (OSU). He received his Ph.D. from Indiana University-Bloomington and is a licensed child/adolescent clinical psychologist, researcher and supervisor.

L. Eugene Arnold is a professor emeritus in the Department of Psychiatry at OSU. He received his MD from OSU and is a child psychiatrist, educator, and a principal investigator in the psychopharmacology and related research program at the OSU Nisonger Center.

Sarah Hersch is a graduate student at Teachers' College, Columbia University, Department of Psychology. She received her BS in business administration from OSU and was a project coordinator on a neurofeedback study at the OSU Nisonger Center.

Elizabeth Hurt is a postdoctoral researcher in the OSU Nisonger Center. She received her PhD from Purdue University and is a clinical site coordinator in the psychopharmacology and related research program at the OSU Nisonger Center.

Roger DeBeus is a clinical assistant professor in the Department of Psychiatry and Behavioral Sciences at East Tennessee State University. He received his PhD from the University of North Texas and is a licensed psychologist and expert, clinician, and researcher in neurofeedback.